

## BOUNDED ARITHMETIC IN FREE LOGIC

YORIYUKI YAMAGATA

National Institute of Advanced Industrial Science and Technology (AIST)  
3-11-46 Nakoji, Amagasaki, 661-0974 Japan  
*e-mail address:* yoriyuki.yamagata@aist.go.jp

**ABSTRACT.** One of the central open questions in bounded arithmetic is whether Buss’ hierarchy of theories of bounded arithmetic collapses or not. In this paper, we reformulate Buss’ theories using *free logic* and conjecture that such theories are easier to handle. To show this, we first prove that Buss’ theories prove consistencies of induction-free fragments of our theories whose formulae have bounded complexity. Next, we prove that although our theories are based on an apparently weaker logic, we can interpret theories in Buss’ hierarchy by our theories using a simple translation. Finally, we investigate finitistic Gödel sentences in our systems in the hope of proving that a theory in a lower level of Buss’ hierarchy cannot prove consistency of induction-free fragments of our theories whose formulae have higher complexity.

### 1. INTRODUCTION

One of the central open questions in bounded arithmetic is whether Buss’ hierarchy  $S_2^1 \subseteq T_2^1 \subseteq S_2^2 \subseteq T_2^2 \subseteq \dots$  of theories of bounded arithmetic collapses [5] or not. Since it is known that collapse of Buss’ hierarchy implies the collapse of the polynomial-time hierarchy [8], demonstration of the non-collapse of the theories in Buss’ hierarchy could be one way to establish the non-collapse of the polynomial-time hierarchy. A natural way to demonstrate non-collapse of the theories in Buss’ hierarchy would be to identify one of these theories that proves (some appropriate formulation of) a statement of the consistency of some theory below it in the hierarchy.

Here, it is clear that we need a delicate notion of consistency because of several negative results that have already been established. The “plain” consistency statement cannot be used to separate the theories in Buss’ hierarchy, since Paris and Wilkie [19] show that  $S_2$  ( $\equiv \bigcup S_2^i$ ) cannot prove the consistency of Robinson Arithmetic  $Q$ . Apparently, this result stems more from the use of predicate logic than from the strength of the base theory. However, Pudlák [14] shows that  $S_2$  cannot prove the consistency of proofs that are carried out within  $S_2^1$  and are comprised entirely of bounded formulae. Even if we restrict our attention to the induction-free fragment of bounded arithmetic, we cannot prove the consistency of such proofs, as shown by Buss and Ignjatović [6]. More precisely, Buss and Ignjatović prove that  $S_2^i$  cannot prove the consistency of proofs that are comprised entirely of  $\Sigma_i^b$  and  $\Pi_i^b$  formulae

*1998 ACM Subject Classification:* F.4.1, F.1.3.

*Key words and phrases:* bounded arithmetic, free logic.

and use only BASIC axioms (the axioms in Buss' hierarchy other than induction) and the rules of inference of predicate logic.

Therefore, if we want to demonstrate non-collapse of the theories in Buss' hierarchy, we should consider a weaker notion of consistency and/or a weaker theory. A number of attempts of this type have been made, both on the positive side (those that establish provability of consistency of some kind) and on the negative side (those that establish non-provability of consistency). On the positive side, Krajíček and Takeuti [9] show that  $T_2^i \vdash \text{RCon}(T_1^i)$ , where  $T_1^i$  is obtained from  $T_2^i$  by eliminating the function symbol  $\#$ , and  $\text{RCon}(T_1^i)$  is a sentence which states that all “regular” proofs carried out within  $T_1^i$  are consistent. Takeuti [17], [18] shows that there is no “small” strictly  $i$ -normal proof  $w$  of contradiction. Here, “ $w$  is small” means that  $w$  has its exponentiation  $2^w$ . Although Takeuti allows induction in strictly  $i$ -normal proof  $w$ , the assumption that  $w$  is small is a significant restriction to  $w$  since bounded arithmetics cannot prove existence of exponentiation. Another direction is to consider cut-free provability. Paris Wilkie [19] mentioned above proves that  $I\Delta_0 + \exp$  proves the consistency of cut-free proofs of  $I\Delta_0$ . For weaker theories than  $I\Delta_0 + \exp$ , we need to relativize the consistency by some cut, then we get similar results [13], [2]. For further weaker theories, Beckmann [4] shows that  $S_2^1$  proves the consistency of  $S_2^{-\infty}$ , where  $S_2^{-\infty}$  is the equational theory which is formalized by recursive definitions of the standard interpretations of the function symbols of  $S_2$ . Also, it is known that  $S_2^i$  proves  $\text{Con}(G_i)$ , that is, the consistency of quantified propositional logic  $G_i$ . On the negative side, we have the results mentioned above, that is, those of Paris and Wilkie [19], Pudlák [14], and Buss and Ignjatović [6]. In addition, there are results which extend incompleteness theorem to Herbrand notion of consistency [3], [1].

In this paper, we introduce the theory  $S_2^i E$  ( $i = -1, 0, 1, 2, \dots$ ), which for  $i \geq 1$  corresponds to Buss'  $S_2^i$ , and we show that the consistency of strictly  $i$ -normal proofs that are carried out only in  $S_2^{-1} E$ , can be proved in  $S_2^{i+2}$ . We improve on the aforementioned positive results in that 1) unlike  $T_1^i$  or  $G_i$ ,  $S_2^i E$  is based on essentially the same language as  $S_2^i$ , thereby making it possible to construct a Gödel sentence by diagonalization; 2) unlike Takeuti [17], [18], we do not assume that the Gödel number of the proofs which are proved consistent are small, that is, have exponentiations, thereby making it possible to apply the second incompleteness theorem—in particular, to derive a Gödel sentence from the consistency statement; 3) unlike the results on Herbrand and cut-free provability,  $S_2^i E$  has the Cut-rule, thereby, making it easy to apply the second incompleteness theorem; and 4) unlike Beckmann [4], our system is formalized in predicate logic. On the other hand, we are still unable to show that the consistency of strictly  $i$ -normal proofs is not provable within  $S_2^j$  for some  $j \leq i$ , but see Section 5. In a sense, our result is an extension of that of Beckmann [4] to predicate logic, since both results are based on the fact that the proofs contain “computations” of the terms that occur in them. In fact, if we drop the Cut-rule from  $S_2^{-1} E$ , the consistency of strictly  $i$ -normal proofs can be proved in  $S_2^1$  for any  $i$ . This “collapse” occurs since, roughly speaking, the combination of the Cut-rule and universal correspond substitution rule in PV.

$S_2^i E$  is based on the following observation: The difficulty in proving the consistency of bounded arithmetic inside  $S_2$  stems from the fact that inside  $S_2$  we cannot define the evaluation function which, given an assignment of natural numbers to the variables, maps the terms of  $S_2$  to their values. For example, the values of the terms  $2, 2\#2, 2\#2\#2, 2\#2\#2\#2, \dots$  increase exponentially; therefore, we cannot define the function that maps these terms to their values, since the rate of growth of every function which is definable in  $S_2$  is dominated

by some polynomial in the length of the input [12]. With a leap of logic, we consider this fact to mean that we cannot assume the existence of values of arbitrary terms in bounded arithmetic. Therefore, we must explicitly prove the existence of values of the terms that occur in any given proof.

Based on this observation,  $S_2^i E$  is formulated by using *free logic* instead of the ordinary predicate calculus. Free logic is a logic which is free from ontological assumptions about the existence of the values of terms. Existence of such objects is explicitly stated by an existential predicate rather than being implicitly assumed. See [11] for a general introduction to free logic and [15] for its application to intuitionistic logic.

Using free logic, we can force each proof carried out within  $S_2^{-1} E$  to somehow “contain” the values of the terms that occur in the proof. By extracting these values from the proof, we can evaluate the terms and then determine the truth value of  $\Sigma_i^b$  formulae. The standard argument using a truth predicate proves the consistency of strictly  $i$ -normal proofs that are carried out only in  $S_2^{-1}$ . It is easy to see that such a consistency proof can be carried out in  $S_2^{i+2}$ .

The paper is organized as follows. In Section 2 we define  $S_2^i E$  and compare it to the systems of logic defined in [11] and [15]. In Section 3 we present the main result of this paper: a proof inside  $S_2^{i+2}$  of the consistency of induction-free strictly  $i$ -normal proofs. In Section 4 we prove the “bootstrapping theorem,” in which we show that although our theory  $S_2^i E$  is based on an apparently weaker logic, we can interpret  $S_2^i$  inside  $S_2^i E$  if  $i \geq 1$ . Unlike the interpretation of  $S_2^i$  by  $Q$  [7], our interpretation does not increase the complexity of formulae; in particular, all bounded formulae are interpreted as bounded formulae. Finally, in Section 5 we raise the question of whether the consistency of induction-free strictly  $i$ -normal proofs can be proved inside  $S_2^i$ . We consider a countably infinite set of finitistic Gödel sentences of  $S_2^{-1} E$  to investigate this question.

## 2. DEFINITION OF $S_2^i E$

$S_2^{-1} E \subseteq S_2^0 E \subseteq S_2^1 E \subseteq S_2^2 E \subseteq \dots$  is a hierarchy of theories resembling Buss’ hierarchy  $S_2^{-1} \subseteq S_2^0 \subseteq S_2^1 \subseteq S_2^2 \subseteq \dots$ . (For purposes of comparison with our system, we include  $S_2^{-1}$  and  $S_2^0$  in Buss’ hierarchy, where  $S_2^{-1}$  is defined as the theory which consists of the formulae that can be proved from the BASIC axioms via the rules of inference of predicate logic, and  $S_2^0$  is the set of formulae that can be proved from the BASIC axioms via the rules of inference of predicate logic together with induction on quantifier-free formulae.) Our system is equipped with a predicate  $E$  which signifies existence of the values of terms. In this section, we introduce the theories  $S_2^i E$  ( $i \geq -1$ ) and their languages, and we prove their basic properties.

**Definition 2.1.** The *theory*  $S_2^i E$  consists of the formulae of  $S_2^i E$  that can be proved from the union of a finite set of logical axioms and a finite set  $\mathcal{A}$  of proper axioms (as defined in Section 2.2) via the rules of inference of free logic (with some modifications) and (for  $i \geq 0$ ) the  $\Sigma_i^b$ -PIND rule; the latter is an induction principle for  $\Sigma_i^b$  formulae which is based on the binary representations of the nonnegative integers. We do not explicitly specify what  $\mathcal{A}$  is. Instead, we make it extensible, and we specify certain conditions that  $\mathcal{A}$  must satisfy. We do this so that the motivation behind the axioms will be more transparent. We also allow the set  $\mathcal{F}$  of function symbols to be extensible.  $S_2^i E(\mathcal{F}, \mathcal{A})$  denotes the individual theory obtained by the function symbols in  $\mathcal{F}$  and the proper axioms in  $\mathcal{A}$ .

An  $S_2^i E$  proof ( $i = -1, 0, 1, \dots$ ) is a formal deduction in which only the rules of inference of  $S_2^i E$  are used. Thus in an  $S_2^i E$  proof, induction is restricted to application of the  $\Sigma_i^b$ -PIND rule; in the  $i = -1$  case, induction is not allowed at all.

The theory  $S_2^i E$  is thus the set of formulae  $A$  of  $S_2^i E$  for which an  $S_2^i E$  proof of the sequent  $\rightarrow A$  exists, which we denote by  $S_2^i E \vdash A$ . We call such formulae *theorems of  $S_2^i E$* . We also use the notation  $\vdash$  for sequents:  $S_2^i E \vdash \Gamma \rightarrow \Delta$ .

In Subsection 2.1 we define the language of  $S_2^i E$ . In Subsection 2.2 we describe the conditions that must be satisfied by  $\mathcal{A}$ , and in Subsection 2.3 we introduce the rules of inference of  $S_2^i E$ . In Subsection 2.4 we compare our system to known systems of free logic.

**2.1. Language of  $S_2^i E$ .** The vocabulary of  $S_2^i E$  is obtained from that of  $S_2^i$  by adding the unary predicate symbol  $E$  and replacing the set of function symbols of  $S_2$  with an arbitrary but finite set  $\mathcal{F}$  of function symbols which denote polynomial-time computable functions. The formulae of  $S_2^i E$  are built up from atomic formulae by use of the propositional connectives  $\neg, \wedge, \vee$ ; the bounded quantifiers  $\forall x \leq t, \exists x \leq t$ ; and the unbounded quantifiers  $\forall x, \exists x$ . Implication ( $\supset$ ) is omitted from the language, and negation ( $\neg$ ) is applied only to equality  $=$  and inequality  $\leq$ . These restrictions appear essential to prove consistency. If there is implication (or negation applied to arbitrary formulae) in  $S_2^{-1} E$ ,  $S_2^{-1} E$  allows induction speedup [6] [14], therefore  $S_2^{-1} E$  polynomially interprets  $S_2^i E$ ,  $i \geq 0$ . This allows to prove  $Ef(n)$  for any polynomial time  $f$  in  $S_2^{-1} E$  by a proof whose length is bounded by some fixed polynomial of length of binary representation of  $n$ . However, this contradicts the statement of soundness (Proposition 3.20)

Since the standard interpretation of all function symbols in  $\mathcal{F}$  are a polynomial-time computable functions, all the function symbols of  $S_2 E$  are definable in  $S_2^1$ , and we assume that Cobham's recursive definitions of polynomial-time computable functions are attached to the corresponding function symbols.

We sometimes identify the function symbols of  $S_2 E$  with their standard interpretations. The distinction between the two types of entities will be clear from the context.

**Definition 2.2.** A set  $\mathcal{F}$  of function symbols (for polynomial-time computable functions) is *well grounded* if it satisfies the following conditions.

- (1)  $\mathcal{F}$  contains the  $n$ -ary constant zero function  $0^n$  (the  $n$ -ary constant function whose value is 0); the  $n$ -ary projection function  $\text{proj}_l^n$  (the  $n$ -ary function that outputs the  $l$ th element in a sequence of length  $n$ ),  $k = 1, \dots, n$ ; and the so-called binary successor functions  $s_0$  and  $s_1$ , where  $s_0$  (resp.  $s_1$ ) is the unary function defined by  $s_0(a) := 2a$  (resp.  $s_1(a) := 2a+1$ ). Note that the binary representation of  $2a$  (resp.  $2a+1$ ) is obtained by appending 0 (resp. 1) to the binary representation of  $a$ , whence the moniker *binary successor function*.
- (2) If  $f \in \mathcal{F}$  is defined from functions  $g, h_1, \dots, h_n$  by composition, then  $g, h_1, \dots, h_n \in \mathcal{F}$ .
- (3) If  $f \in \mathcal{F}$  is defined from functions  $g, h_1, h_2$  by recursion, that is,  $f$  is defined by the equations

$$f(0, \vec{x}) = g(\vec{x}) \tag{2.1}$$

$$f(s_0 x, \vec{x}) = h_0(x, \vec{x}, f(x, \vec{x})) \tag{2.2}$$

$$f(s_1 x, \vec{x}) = h_1(x, \vec{x}, f(x, \vec{x})) \tag{2.3}$$

then  $g, h_1, h_2 \in \mathcal{F}$ . Cobham's limit recursion on notations can be written as above, providing  $\text{Cond}, \# \in \mathcal{F}$ . Since  $\text{Cond}, \#$  can be defined by recursion as above, for any finite set of polynomial time functions  $f_1, \dots, f_n$ , there is a well-grounded  $\mathcal{F}$  such that  $f_1, \dots, f_n \in \mathcal{F}$ .

If  $\mathcal{F}$  is well grounded (which is true of every set  $\mathcal{F}$  of function symbols we consider in this paper), we can define the *definition degree*  $d(f)$  of each  $f \in \mathcal{F}$ .

**Definition 2.3** (Definition degree). Let  $f \in \mathcal{F}$ .

- (1) If  $f$  is  $0^n$ ,  $\text{proj}_i^n$ ,  $s_0$ ,  $s_1$ , then  $d(f) := 0$ .
- (2) If  $f$  is defined from  $g, h_1, \dots, h_n$  by composition, then

$$d(f) := 1 + \max\{d(g), d(h_1), \dots, d(h_n)\} \quad (2.4)$$

- (3) If  $f$  is defined from  $g, h_1, h_2$  by recursion, then

$$d(f) := 1 + \max\{d(g), d(h_1), d(h_2)\} \quad (2.5)$$

Induction on the definition degree of  $f$  is used in Subsection 4.2 to prove the totality of  $f$  (i.e., that  $Ea_1, \dots, Ea_n \rightarrow Ef(a_1, \dots, a_n)$ , where  $a_1, \dots, a_n$  are used here as meta-symbols for variables of  $S_2^i E$ ). Now we define the vocabulary of  $S_2^i E(\mathcal{F}, \mathcal{A})$ .

**Definition 2.4** (Vocabulary). The *vocabulary* of  $S_2^i E(\mathcal{F}, \mathcal{A})$  consists of the following symbols.

**Constant symbols:** The only constant symbol of  $S_2^i$  is 0.

**Variables:** The variables of  $S_2^i E$  are  $x_1, x_2, \dots$ . We often use  $a, b$  and  $a_1, a_2, \dots, b_1, b_2, \dots, x, y, x_1, x_2, \dots$  as meta-symbols for variables, and we often denote a finite sequence of variables by  $\vec{a}$ ,  $\vec{a}'$ , or  $\vec{b}$ .

**Function symbols:** The function symbols of  $S_2^i E(\mathcal{F}, \mathcal{A})$  are the symbols in the finite set  $\mathcal{F}$ . For all  $i \geq 0$ , we can interpret  $S_2^i$  in  $S_2^i E(\mathcal{F}, \mathcal{A})$ , provided that  $\mathcal{F}$  contains the function symbols of  $S_2$  (those for the unary functions  $S$ ,  $\lfloor \frac{\cdot}{2} \rfloor$ ,  $|\cdot|$  and the binary functions  $\#, +, \cdot$ ), where  $S$  is the successor function division by two rounded to 0, length defined by  $S(a) := a + 1$  and  $\lfloor \frac{\cdot}{2} \rfloor$  is the function where  $\lfloor \frac{a}{2} \rfloor$  is defined as the natural number  $n$  such that  $a \in \{2n, 2n + 1\}$ . As stated earlier,  $|a|$  is defined as the number of bits in the binary representation of  $a$  (by convention,  $|0| = 0$ ). The binary function  $\#$  is the so-called smash function defined by  $a \# b := 2^{|a||b|}$ , and  $+$  and  $\cdot$  are the usual addition and multiplication functions, respectively. We assume that these function symbols are contained in  $\mathcal{F}$ . Further as we exploit the binary representations of the natural numbers (as finite bit strings), we introduce the binary function  $\oplus$  and  $\ominus$ .  $\oplus(a, b)$  is defined as the natural number whose binary representation is the concatenation of the binary representations of the natural numbers  $a$  and  $b$  (the bits of  $a$  being the most significant bits of  $\oplus(a, b)$ ).  $\ominus(a, b)$  is defined as  $\lfloor \frac{a}{2^{|b|}} \rfloor$ . We also assume  $\oplus$  and  $\ominus$  are contained in  $\mathcal{F}$ .

From this point on, we write them in infix notation (as in  $u \oplus r$  and  $u \ominus r$ ). Since the functions  $\ominus$  and  $\oplus$  are polynomial-time functions, and thus  $\Sigma_1^b$ -definable in  $S_2^1$ , we use the notations  $u \oplus r$  and  $u \ominus r$  in our informal proofs as well as in actual formulae of  $S_2$  and  $S_2 E$ .

**Predicate symbols:**  $S_2^i E$  has three predicate symbols:  $E$ ,  $=$ ,  $\leq$ . The unary predicate  $E$  signifies that, for every term  $t$  of which  $E$  is asserted to hold, the value of  $t$  actually exists (i.e., that it converges to a standard natural number). The binary predicate  $=$  signifies equality, and  $\leq$  signifies the less-than-or-equal-to relation.  $p$  is used as a meta-variable for the predicate symbols  $=$  and  $\leq$ .

**Definition 2.5** (Terms). The *terms* of  $S_2^i E$  are defined recursively as follows.

- 0 and the variables  $x_1, x_2, x_3, \dots$  are terms.
- If  $f$  is an  $n$ -ary function symbol and  $t_1, \dots, t_n$  are terms, then  $f(t_1, \dots, t_n)$  is a term.

We use  $s, t, t_1, t_2, \dots, u_1, u_2, \dots$  as meta-variables for terms, and we often denote a finite sequence of terms by  $\vec{s}$  or  $\vec{t}$ .

For the functions  $S, s_0, s_1$ , we omit parentheses, instead denoting  $S(t)$ ,  $s_0(t)$ , and  $s_1(t)$  by  $St$ ,  $s_0t$ , and  $s_1t$ , respectively. Also, we write  $\lfloor \frac{t}{2} \rfloor$  or  $\lfloor \frac{1}{2}t \rfloor$  for  $\lfloor \frac{\cdot}{2} \rfloor t$ , and we write the binary functions  $\#, +, \cdot$  in infix notation, as in  $t + u$ .

We define *numerals* as terms which are constructed from  $s_0$ ,  $s_1$ , and 0 alone. They are used in  $S_2^i E$  to denote natural numbers. We (informally) use the numbers  $0, 1, 2, \dots$  (represented in the decimal system) to denote numerals. In general, we use the same notation for numerals as for the corresponding natural numbers. The distinction between the two types of entities will be obvious from the context.

**Definition 2.6** (Formulae). The *formulae* of  $S_2^i E$  are defined recursively as follows.

- If  $t$  is a term, then  $Et$  is a formula.
- If  $p$  is  $=$  or  $\leq$  and  $t_1, t_2$  are terms, then  $p(t_1, t_2)$  is a formula.
- If  $p$  is  $=$  or  $\leq$  and  $t_1, t_2$  are terms, then  $\neg p(t_1, t_2)$  is a formula.
- If  $A$  and  $B$  are formulae, then  $A \wedge B$  and  $A \vee B$  are formulae.
- If  $A(a)$  is a formula and  $x$  is a variable, then  $\forall x.A(x)$  and  $\exists x.A(x)$  are formulae.
- If  $A(a)$  is a formula,  $x$  is a variable, and  $t$  is a term, then  $\forall x \leq t.A(x)$  and  $\exists x \leq t.A(x)$  are formulae.

We use  $A, B, \dots, A_0, A_1, A_2, \dots$  as meta-variables for formulae, and  $\Gamma, \Delta, \Pi, \Lambda$ , and  $\Gamma_1, \Gamma_2, \dots, \Delta_1, \Delta_2, \dots$  as meta-variables for finite sequences of formulae.  $\Gamma, \Pi$  denotes the concatenation of  $\Gamma$  and  $\Pi$  (in that order:  $\Gamma$  followed by  $\Pi$ ). The concatenation of a finite sequence  $\Gamma$  and a single formula  $A$  is denoted by  $\Gamma, A$ , while the concatenation of  $A$  and  $\Gamma$  is denoted by  $A, \Gamma$ .

If  $p$  is  $=$  (resp.  $\leq$ ), we write the formula  $p(t_1, t_2)$  in infix notation, as  $t_1 = t_2$  (resp.  $t_1 \leq t_2$ ). We write the negations of  $t_1 = t_2$  and  $t_1 \leq t_2$  as  $t_1 \neq t_2$  and  $t_1 \not\leq t_2$ , respectively. In the meta-language, a notation of the form  $a \equiv b$  means that  $a$  and  $b$  are the same syntactic construct (the same variable, the same term, the same formula, or the same sequence of variables/terms). We often write  $E\vec{a}$  to denote the sequence  $Ea_1, \dots, Ea_k$  for a finite sequence  $\vec{a} \equiv a_1, \dots, a_k$  of variables.

Formulae have the usual meaning, and free and bound variables are defined as usual. We sometimes write  $A(a)$  to indicate that  $a$  is possibly among the variables that occur free in  $A$ . This notation does not imply that  $a$  actually occurs in  $A$ ; what it indicates is that if  $a$  does occur in  $A$ , then  $a$  occurs free in  $A$ . We tend to use meta-symbols such as  $x, y, x_1, x_2, \dots$  for variables that are captured by an outer quantifier, as in  $\forall x.A(x)$  or  $\exists x.A(x)$ , and meta-symbols such as  $a, b, a_1, a_2, \dots, b_1, b_2, \dots$  otherwise. We write  $A(t)$  for the formula which is obtained by substituting the term  $t$  for  $a$  in  $A(a)$  (or for the variable  $x$  in  $A(x)$ ). We follow the convention of assuming that variables which occur bound in  $A(a)$  and also occur in  $t$  are renamed in  $A(a)$  before such a substitution is made, so that the variables in  $t$  are not “accidentally” bound in  $A(t)$ .

We say that a quantifier is *bounded* if it is of the form  $\forall x \leq t$  or  $\exists x \leq t$ . On the other hand, the quantifiers  $\forall x$  and  $\exists x$ , without the bound  $\leq t$ , are called *unbounded*.

A formula of the form  $t_1 = t_2$  or  $t_1 \leq t_2$  is called *atomic*, and a formula of the form  $Et$ ,  $t_1 = t_2$ ,  $t_1 \leq t_2$ ,  $t_1 \neq t_2$ , or  $t_1 \not\leq t_2$  is called *basic*. Also, a formula of the form  $Et$  is called

an  $E$ -form, but in contrast to the usual convention for atomic formulae in predicate logic, an  $E$ -form is defined not to be atomic.

In  $S_2^i E$ , implication is absent and negation is restricted to atomic formulae. We interpret the negation  $\neg A$  for a non  $E$ -form formula  $A$  as the De Morgan dual  $\overline{A}$ , and the implication  $A \subset B$  as  $\overline{A} \vee B$ . As shown in Section 4.4,  $\neg$  and  $\supset$  defined in this way satisfy the usual rules of inference of predicate logic, provided that all the functions represented by function symbols of  $S_2 E$  are provably total (that is, for each  $n$ -ary function symbol  $f$ , we can prove  $Ea_1, \dots, Ea_n \rightarrow Ef(a_1, \dots, a_n)$ ). This fact is proved in Proposition 4.3.

Since implication is absent, we cannot code bounded universal quantification by unbounded universal quantification together with implication. Therefore, we need bounded quantifiers that are constructs in their own right.

By formalizing the notion of “usual meaning,” we obtain an interpretation of the formulae of  $S_2^i E$  in  $S_2$ .

**Definition 2.7** (Standard Interpretation). We interpret an  $S_2^i E$  formula  $A$  in  $S_2$  by first replacing every subformula of  $A$  which is of the form  $Et$  with the  $S_2$  formula  $t = t$ , and then replacing the function symbols in the resulting formula with their  $S_2^1$  definitions. Therefore, we obtain an interpretation of the  $S_2^i E$  formulae in  $S_2^i$ , and for  $i \in \{-1, 0\}$  we obtain an interpretation of the  $S_2^i E$  formulae in  $S_2^1$ .

In Subsection 4.1, we present an interpretation of the  $S_2^i$  formulae in  $S_2^i E(\mathcal{F}, \mathcal{A})$  for any  $\mathcal{F}$  that contains the function symbols of  $S_2$ .

We say that a formula of  $S_2^i E$  is *bounded* if all of its quantifiers are bounded. The bounded formulae of  $S_2^i E$  are classified into hierarchies  $\Sigma_j^b, \Pi_j^b$  ( $j \in \mathbb{N}$ ).

**Definition 2.8** ( $\Sigma_j^b, \Pi_j^b$ ). The classes  $\Sigma_j^b, \Pi_j^b$  of bounded formulae of  $S_2^i E$  are defined recursively as follows.

- $\Sigma_0^b = \Pi_0^b$  is the set of quantifier-free formulae (formulae without quantifiers).
- If  $A, B \in \Sigma_j^b$ , then  $A \wedge B, A \vee B \in \Sigma_j^b$ . Similarly, if  $A, B \in \Pi_j^b$ , then  $A \wedge B, A \vee B \in \Pi_j^b$ .
- If  $A \in \Sigma_j^b$ , then  $A \in \Pi_{j+1}^b$ . Similarly, if  $A \in \Pi_j^b$ , then  $A \in \Sigma_{j+1}^b$ .
- If  $A(a) \in \Sigma_j^b$ , then  $\forall x \leq t. A(x) \in \Pi_{j+1}^b$ . Similarly, if  $A(a) \in \Pi_j^b$ , then  $\exists x \leq t. A(x) \in \Sigma_{j+1}^b$ .

These hierarchies are used to control the strength of mathematical induction in  $S_2 E$ , just as in Buss' system.

**2.2. Axioms of  $S_2^i E$ .** In this subsection, we first discuss the conditions which all axioms of  $S_2^i E$  must satisfy. Then we introduce the logical axioms, and finally we impose certain conditions on the axioms in  $\mathcal{A}$  that will allow us to interpret  $S_2^i$  inside  $S_2^i E$  for  $i \geq 1$ .

For our proof of consistency of strictly  $i$ -normal proofs to work, the axioms of  $S_2^i E$  must satisfy *the boundedness conditions*.

**Definition 2.9.** A sequent  $\Gamma \rightarrow \Delta$  *satisfies the boundedness conditions* if it has the following three properties, where  $\vec{a}$  are the variables that occur free in  $\Gamma \rightarrow \Delta$ .

- (1) All the formulae that occur in  $\Gamma \rightarrow \Delta$  are basic.
- (2) Every variable in  $\vec{a}$  occurs free in at least one formula in  $\Gamma$ .
- (3) There is a constant  $\alpha \in \mathbb{N}$  such that

$$S_2^1 \vdash \max\{\vec{t}_\Delta(\vec{a})\} \leq \alpha \cdot \max\{\vec{t}_\Gamma(\vec{a})\}, \quad (2.6)$$

$\vec{t}_\Delta(\vec{a})$  are the subterms of the terms that occur in  $\Delta$  and  $\vec{t}_\Gamma(\vec{a})$  are the subterms of the terms that occur in  $\Gamma$  (for convenience,  $\max \emptyset$  is defined to be 1). Since the function symbols of  $S_2E$  are definable in  $S_2^1$ , we can regard the terms in  $\vec{t}_\Gamma(\vec{a})$  and  $\vec{t}_\Delta(\vec{a})$  as terms of  $S_2^1$ , hence we can regard  $\max\{\vec{t}_\Gamma(\vec{a})\} \leq \alpha \cdot \max\{\vec{t}_\Delta(\vec{a})\}$  as an  $S_2^1$  formula.

For our proof of consistency of  $i$ -normal proofs, we need the boundedness conditions to hold of the axioms. The third condition states that, under any given valuation of the variables that occur free in  $\Gamma \rightarrow \Delta$ , the values of the subterms of the terms in  $\Delta$  cannot exceed the values of the subterms of the terms in  $\Gamma$  by the constant factor  $\alpha$ , and that this fact can be proved in  $S_2^1$ . In general, the constant  $\alpha$  varies from one axiom to another. Since  $\mathcal{A}$  is finite, however, there is a single constant  $\alpha$  that applies to all the axioms in  $\mathcal{A}$ . For the logical axioms and the proper axioms required to interpret  $S_2^iE$  in  $S_2^1$ , we can take  $\alpha$  to be 4, so from this point on we assume that  $\alpha = 4$ . This property plays a crucial role in our consistency proof. The reason for imposing the first and second boundedness conditions is that we want to avoid complexities that would otherwise arise in satisfying the third boundedness condition.

The logical axioms are divided into *E-axioms* and *equality axioms*. The *E-axioms* establish the basic properties of the  $E$  predicate, and the equality axioms establish the basic properties of equality.

**Definition 2.10** (*E-axioms*).

$$Ef(a_1, \dots, a_n) \rightarrow Ea_j \quad j = 1, \dots, n; f \in \mathcal{F} \quad (2.7)$$

$$p(a_1, a_2) \rightarrow Ea_j \quad j = 1, 2; p \in \{=, \leq\} \quad (2.8)$$

$$\neg p(a_1, a_2) \rightarrow Ea_j \quad j = 1, 2; p \in \{=, \leq\} \quad (2.9)$$

Axiom (2.7) states that if the value of  $f$  exists, then the arguments of  $f$  also exist. Axioms (2.8) and (2.9) state that if  $=$ ,  $\leq$ , or the negation of either of them holds of  $a_1, a_2$ , then  $a_1, a_2$  exist. That is, in the terminology of Scott [15], we assume that all functions and predicates are *strict*.

**Definition 2.11** (*Equality axioms*).

$$Ea \rightarrow a = a \quad (2.10)$$

$$a = b \rightarrow b = a \quad (2.11)$$

$$a = b, b = c \rightarrow a = c \quad (2.12)$$

$$a = b \rightarrow s_j a = s_j b \quad j = 0, 1 \quad (2.13)$$

$$Ef(\vec{a}), \vec{a} = \vec{b} \rightarrow f(\vec{a}) = f(\vec{b}) \quad f \in \mathcal{F} \quad (2.14)$$

$$a_1 = b_1, a_2 = b_2, p(a_1, a_2) \rightarrow p(b_1, b_2) \quad p \in \{=, \leq\} \quad (2.15)$$

Axioms (2.11), (2.12), (2.13), and (2.15) are standard, but Axiom (2.10) needs an explanation. For our soundness proof, we would like to have the property that when a closed formula  $\phi$  is proved, the values of the terms which occur in  $\phi$  are bounded by the code for the proof of  $\phi$ . Therefore, we cannot use just  $t = t$  as a substitution instance of an equality axiom with a closed term  $t$ , since we would have no control over the value of  $t$ . To deal with this problem, we add  $Ea$  to the antecedent of the usual reflexive law of equality, which ensures that the proof of  $t = t$  contains the proof of  $Et$ , and hence that the code for the proof of  $t = t$  exceeds the value of  $t$ .



Axiom (2.13) states that the binary successor functions  $s_0, s_1$  preserve equality. For the other function symbols, we have only a restricted preservation of equality,  $Ef(\vec{a}), \vec{a} = \vec{b} \rightarrow f(\vec{a}) = f(\vec{b})$ , as an axiom, and not full preservation of equality,  $\vec{a} = \vec{b} \rightarrow f(\vec{a}) = f(\vec{b})$ . This is because, the case of the reflexive law of equality, for sequences  $\vec{t}, \vec{u}$  of closed terms we want to bound the values of  $f(\vec{t})$  and  $f(\vec{u})$  by the code for the proof of the equality  $f(\vec{t}) = f(\vec{u})$ . Without the condition  $Ef(\vec{a})$  in the antecedent of Axiom (2.14), we cannot ensure that this property will hold. For example, if  $t_1, t_2, u_1, u_2$  are closed terms and  $t_1 = u_1, t_2 = u_2 \rightarrow t_1 \# t_2 = u_1 \# u_2$  is a substitution instance of the usual equality axiom (Axiom (2.14) with  $f \equiv \#, \vec{a} \equiv (a_1, a_2)$ , and  $\vec{b} \equiv (b_1, b_2)$ , but without  $Ef(\vec{a})$  in the antecedent), the values of  $t_1 \# t_2$  and  $u_1 \# u_2$  are not necessarily bounded by the code for the proof of the equality  $t_1 \# t_2 = u_1 \# u_2$ . However, we can prove full preservation of equality for function symbols as a theorem, without assuming  $Ef(\vec{a})$  (see Subsection 4.3), since we can prove that all the function symbols of  $S_2E$  represent total functions, that is, that for all  $f \in \mathcal{F}$ ,  $E\vec{a} \rightarrow Ef(\vec{a})$ . Actually, using the  $\Sigma_2^b$ -PIND rule, we can prove full preservation of equality for function symbols by induction on the definition degree of  $f$ , without using Axiom (2.14) at all. However, we retain Axiom (2.14) since we want to have the equality axioms in  $S_2^1E$ .

Axiom (2.15) states that the truth value of the predicates  $=$  and  $\leq$  is preserved by equality.

Next, we discuss the proper axioms (the axioms in  $\mathcal{A}$ ). In  $S_2^{i+2}$  we can prove the consistency of strictly  $i$ -normal proofs—which are carried out within  $S_2^{-1}E(\mathcal{F}, \mathcal{A})$ —for any  $\mathcal{A}$  all of whose elements satisfy the boundedness conditions. To interpret  $S_2^i$  inside  $S_2^iE(\mathcal{F}, \mathcal{A})$ , we need to have the function symbols of  $S_2$  in  $\mathcal{F}$ , and to have the following data axioms, separation axioms and auxiliary axioms in  $\mathcal{A}$ .

**Definition 2.12** (Data axioms). These are the *data axioms* introduced by Leivant [10].

$$\rightarrow E0 \quad (2.16)$$

$$Ea \rightarrow Es_j a \quad j = 0, 1 \quad (2.17)$$

**Definition 2.13** (Separation axioms).

$$s_j a = s_j b \rightarrow a = b \quad j = 0, 1 \quad (2.18)$$

$$Ea, Eb \rightarrow s_0 a \neq s_1 b \quad (2.19)$$

$$Ea \rightarrow s_1 a \neq 0 \quad (2.20)$$

$$s_0 a = 0 \rightarrow a = 0 \quad (2.21)$$

**Definition 2.14** (Defining axioms). It is clear that by Cobham's definition of polynomial-time computable functions, the defining equations for each  $n$ -ary function symbol  $f \in \mathcal{F}$  can be written in the form  $f(u(a_1), a_2, \dots, a_n) = t(a_1, a_2, \dots, a_n)$  where  $u(a_1)$  is one of  $0, a_1, s_0 a_1, s_1 a_1$ . For each defining equation in this form, the following axiom is in  $\mathcal{A}$ .

$$Ea_1, Ea_2, \dots, Ea_n, Et(a_1, a_2, \dots, a_n) \rightarrow f(u(a_1), a_2, \dots, a_n) = t(a_1, a_2, \dots, a_n) \quad (2.22)$$

We call these axioms the *defining axioms*.

**Definition 2.15** (Auxiliary axioms). The *auxiliary axioms* are those generated from the BASIC axioms of  $S_2$  by the following procedure.

First, render these axioms free of logical connectives (see [17]). For example, the first BASIC axiom,  $b \leq a \supset b \leq Sa$ , is changed to the sequent  $b \leq a \rightarrow b \leq Sa$ .

Then for each term  $t$  that occurs in the succedent, we add  $Et$  to the antecedent. The axiom in the example becomes  $Eb, ESa, b \leq a \rightarrow b \leq Sa$ .

Finally, simplify the resulting axioms by removing the unnecessary assumptions—taking care, however, to ensure that the final version of each axiom satisfies the boundedness conditions (Definition 2.9). For example, we simplify  $Eb, ESa, b \leq a \rightarrow b \leq Sa$  to  $ESa, b \leq a \rightarrow b \leq Sa$ , since  $Eb$  follows from  $b \leq a$  by  $E$ -axiom (2.8) and, after this simplification, this auxiliary axiom still satisfies the boundedness conditions.

From the data axioms and separation axioms, we can prove the relations  $n = m$  and  $n \neq m$  that hold between numerals  $n, m$  in their standard interpretations. Moreover, we can prove totality of every  $f \in \mathcal{F}$  (i.e., that  $Ea_1, \dots, Ea_n \rightarrow Ef(a_1, \dots, a_n)$ ) from the logical axioms, the defining axioms, and  $\Sigma_0^b$ -PIND (see Subsection 4.2). Also, we can use the auxiliary axioms, together with other axioms and the rules of inference of  $S_2^1E$ , to prove the interpretation of the BASIC axioms of  $S_2$  in  $S_2^1E$  which is defined in Subsection 4.1 (see Subsection 4.4).

From this point on, we assume that  $\mathcal{F}$  contains all the function symbols of  $S_2$  and that  $\mathcal{A}$  contains all the proper axioms outlined above. All of these proper axioms, as well as the logical axioms, satisfy the boundedness conditions.

**2.3. Rules of inference of  $S_2^iE$ .** We formulate  $S_2^iE$  proofs using the sequent calculus. The rules of inference of  $S_2^iE$  consist of the rules for predicate logic (but with a modified  $R \neg$ -rule; see (2.33)) plus the PIND rule applied to  $\Sigma_i^b$  formulae, which is called the  $\Sigma_i^b$ -PIND rule. The  $\Sigma_i^b$ -PIND rule derives  $\Gamma, Et \rightarrow A(t), \Delta$  from  $\Gamma \rightarrow \Delta, A(0); A(a), \Gamma \rightarrow \Delta, A(s_0a)$ ; and  $A(a), \Gamma \rightarrow \Delta, A(s_1a)$ .

If we can derive a sequent  $\Gamma \rightarrow \Delta$  from premises  $\Gamma_1 \rightarrow \Delta_1, \dots, \Gamma_n \rightarrow \Delta_n$ , we write

$$\frac{\Gamma_1 \rightarrow \Delta_1 \quad \cdots \quad \Gamma_n \rightarrow \Delta_n}{\Gamma \rightarrow \Delta} R, \quad (2.23)$$

where  $R$  is the name of the rule. The name is often omitted if it is obvious. All the rules of inference which appear in this paper have at most three premises.

The rules of inference are divided into the *identity rule*, the *axiom rule*, the *structural rules*, the *logical rules*, the *Cut rule*, and the *PIND rule*.

The *identity rule* is used to express that a formula  $A$  implies itself.

Identity rule:

$$\overline{A \rightarrow A} \text{ Id} \quad (2.24)$$

The *axiom rule* is used to derive a sequent which is a *substitution instance* of an axiom, that is, a sequent of the form  $\Gamma(\vec{t}(\vec{b})) \rightarrow \Delta(\vec{t}(\vec{b}))$  where  $\Gamma(\vec{a}) \rightarrow \Delta(\vec{a})$  is an axiom,  $\vec{a}$  are the variables that occur in  $\Gamma \rightarrow \Delta$ ,  $\vec{t}$  are the terms that are being substituted for the variables in  $\vec{a}$ , and  $\vec{b}$  are the variables that occur in  $\vec{t}$ .

Axiom rule:

$$\overline{\Gamma(\vec{t}(\vec{b})) \rightarrow \Delta(\vec{t}(\vec{b}))} \text{ Ax} \quad (2.25)$$

The *structural* rules are defined as usual.

Weakening rules:

$$\frac{\Gamma \rightarrow \Delta}{A, \Gamma \rightarrow \Delta} \quad (2.26)$$

$$\frac{\Gamma \rightarrow \Delta}{\Gamma \rightarrow \Delta, A} \quad (2.27)$$

Contraction rules:

$$\frac{A, A, \Gamma \rightarrow \Delta}{A, \Gamma \rightarrow \Delta} \quad (2.28)$$

$$\frac{\Gamma \rightarrow \Delta, A, A}{\Gamma \rightarrow \Delta, A} \quad (2.29)$$

Exchange rules:

$$\frac{\Gamma, A, B, \Pi \rightarrow \Delta}{\Gamma, B, A, \Pi \rightarrow \Delta} \quad (2.30)$$

$$\frac{\Gamma \rightarrow \Delta, A, B, \Pi}{\Gamma \rightarrow \Delta, B, A, \Pi} \quad (2.31)$$

The logical rules are those of classical sequent calculus with a modified  $R\neg$ -rule. The logical rules for negation are used to negate atomic formulae only, since the language of  $S_2^iE$  allows negation of atomic formulae only.

$\neg$ -rules:

$$\frac{\Gamma \rightarrow \Delta, p(t_1, t_2)}{\neg p(t_1, t_2), \Gamma \rightarrow \Delta} \text{L} \neg \quad (2.32)$$

$$\frac{p(t_1, t_2), \Gamma \rightarrow \Delta}{Et_1, Et_2, \Gamma \rightarrow \Delta, \neg p(t_1, t_2)} \text{R} \neg, \quad (2.33)$$

where  $p$  is  $=$  or  $\leq$  and  $t_1, t_2$  are terms. Unlike the usual textbook definition of the  $R\neg$ -rule, in the antecedent we introduce the formulae  $Et_1$  and  $Et_2$  that express the existence of values of the terms  $t_1$  and  $t_2$ , respectively. This is because we interpret  $\neg p(t_1, t_2)$  to mean that the values  $v_1, v_2$  of  $t_1, t_2$  exist and satisfy  $\neg p(v_1, v_2)$ . Therefore, to infer  $\neg p(t_1, t_2)$  we presuppose  $Et_1$  and  $Et_2$ .

$\wedge$ -rules:

$$\frac{A, \Gamma \rightarrow \Delta}{A \wedge B, \Gamma \rightarrow \Delta} \text{L} \wedge_1 \quad (2.34)$$

$$\frac{A, \Gamma \rightarrow \Delta}{B \wedge A, \Gamma \rightarrow \Delta} \text{L} \wedge_2 \quad (2.35)$$

$$\frac{\Gamma \rightarrow \Delta, A \quad \Gamma \rightarrow \Delta, B}{\Gamma \rightarrow \Delta, A \wedge B} \text{R} \wedge \quad (2.36)$$

$\vee$ -rules:

$$\frac{A, \Gamma \rightarrow \Delta \quad B, \Gamma \rightarrow \Delta}{A \vee B, \Gamma \rightarrow \Delta} \text{L}\vee \quad (2.37)$$

$$\frac{\Gamma \rightarrow \Delta, A}{\Gamma \rightarrow \Delta, A \vee B} \text{R}\vee_1 \quad (2.38)$$

$$\frac{\Gamma \rightarrow \Delta, A}{\Gamma \rightarrow \Delta, B \vee A} \text{R}\vee_2 \quad (2.39)$$

The logical rules for universal quantification are of two types: *bounded* (indicated with a “b” following the L or R) and *unbounded*, corresponding to bounded and unbounded universal quantification, respectively, in the language of  $S_2^i E$ . The same is true of the logical rules for existential quantification. In both types, bounded and unbounded, the quantification (universal or existential) is done over objects that actually exist.

Bounded  $\forall$ -rules:

$$\frac{A(t), \Gamma \rightarrow \Delta}{t \leq s, \forall x \leq s. A(x), \Gamma \rightarrow \Delta} \text{Lb}\forall, \quad (2.40)$$

where the variable  $x$  does not occur in the term  $s$ .

$$\frac{a \leq t, \Gamma \rightarrow \Delta, A(a)}{Et, \Gamma \rightarrow \Delta, \forall x \leq t. A(x)} \text{Rb}\forall, \quad (2.41)$$

where neither the variable  $a$  nor the variable  $x$  occurs in the term  $t$ , and  $a$  does not occur free in  $\Gamma \rightarrow \Delta$ .

Unbounded  $\forall$ -rules:

$$\frac{A(t), \Gamma \rightarrow \Delta}{Et, \forall x. A(x), \Gamma \rightarrow \Delta} \text{L}\forall \quad (2.42)$$

$$\frac{Ea, \Gamma \rightarrow \Delta, A(a)}{\Gamma \rightarrow \Delta, \forall x. A(x)} \text{R}\forall, \quad (2.43)$$

where the variable  $a$  does not occur free in  $\Gamma \rightarrow \Delta$ .

Bounded  $\exists$ -rules:

$$\frac{a \leq t, A(a), \Gamma \rightarrow \Delta}{\exists x \leq t. A(x), \Gamma \rightarrow \Delta} \text{Lb}\exists, \quad (2.44)$$

where neither the variable  $a$  nor the variable  $x$  occurs in neither the term  $t$ , and  $a$  does not occur free in  $\Gamma \rightarrow \Delta$ .

$$\frac{\Gamma \rightarrow \Delta, A(t)}{t \leq s, \Gamma \rightarrow \Delta, \exists x \leq s. A(x)} \text{Rb}\exists, \quad (2.45)$$

where the variable  $x$  does not occur in the term  $s$ .

Unbounded  $\exists$ -rules:

$$\frac{Ea, A(a), \Gamma \rightarrow \Delta}{\exists x. A(x), \Gamma \rightarrow \Delta} \text{L}\exists, \quad (2.46)$$

where the variable  $a$  does not occur free in  $\Gamma \rightarrow \Delta$ .

$$\frac{\Gamma \rightarrow \Delta, A(t)}{Et, \Gamma \rightarrow \Delta, \exists x. A(x)} \text{R}\exists \quad (2.47)$$

The *Cut rule* is used to derive a sequent by employing an intermediate “lemma.”

Cut rule:

$$\frac{\Gamma \rightarrow \Delta, A \quad A, \Pi \rightarrow \Lambda}{\Gamma, \Pi \rightarrow \Delta, \Lambda} \text{ Cut} \quad (2.48)$$

The PIND *rule* is used to infer statements  $A(t)$  (for formulae  $A(a)$  and terms  $t$ ) from the assumption that  $t$  converges and  $A$  satisfies the induction hypothesis on the value of  $t$ .

PIND rule:

$$\frac{\Gamma \rightarrow \Delta, A(0) \quad A(a), \Gamma \rightarrow \Delta, A(s_0a) \quad A(a), \Gamma \rightarrow \Delta, A(s_1a)}{Et, \Gamma \rightarrow \Delta, A(t)} \text{ PIND}, \quad (2.49)$$

where the variable  $a$  does not occur free in  $\Gamma \rightarrow \Delta$ . We call this the PIND rule on  $t$ . If we restrict the PIND rule to formulae  $A(a)$  from some class  $\Phi$ , we call the resulting rule the  $\Phi$ -PIND rule.

Informally, the PIND rule expresses induction on the notations of the natural numbers represented in binary. We call  $A(a)$  the *induction hypothesis*, the proof of  $\Gamma \rightarrow \Delta, A(0)$  the *base case*, and the proofs of  $A(a), \Gamma \rightarrow \Delta, A(s_0a)$  and  $A(a), \Gamma \rightarrow \Delta, A(s_1a)$  the *induction steps*.

From the PIND rule on  $t$ , we have the following derived rule.

$$\frac{\Gamma \rightarrow \Delta, A(0) \quad Ea, A(a), \Gamma \rightarrow \Delta, A(s_0a) \quad Ea, A(a), \Gamma \rightarrow \Delta, A(s_1a)}{Et, \Gamma \rightarrow \Delta, A(t)} \quad (2.50)$$

We call this the PIND- $E$  rule on  $t$ . To derive this rule, we first derive  $A(a) \wedge Ea, \Gamma \rightarrow \Delta, A(s_0a) \wedge Es_0a$  from  $Ea, A(a), \Gamma \rightarrow \Delta, A(s_0a)$ . This is done via the axiom  $Ea \rightarrow Es_0a$  and propositional reasoning. Similarly, we derive  $A(a) \wedge Ea, \Gamma \rightarrow \Delta, A(s_1a) \wedge Es_1a$  from  $Ea, A(a), \Gamma \rightarrow \Delta, A(s_1a)$ . Then, applying PIND to  $A(a) \wedge Ea$ , we derive  $Et, \Gamma \rightarrow \Delta, A(t) \wedge Et$ . Finally, we derive  $Et, \Gamma \rightarrow \Delta, A(t)$  by purely propositional reasoning.

This completes the definition of the rules of inference of  $S_2^iE$ .

**Lemma 2.16** (Substitution Lemma). *If  $S_2^iE \vdash \Gamma(a) \rightarrow \Delta(a)$ , where the variable  $a$  occurs free in  $\Gamma \rightarrow \Delta$ , then  $S_2^iE \vdash \Gamma(t) \rightarrow \Delta(t)$  for every term  $t$ .*

*Proof.* Induction on the structure of the  $S_2^iE$  proof of  $\Gamma(a) \rightarrow \Delta(a)$ . □

**2.4. Comparison to Free Logic and Scott's system.** In this subsection, we briefly discuss the similarities between our system and two others: free logic (see [11] for a comprehensive summary) and Scott's  $E$ -logic [15].

Our system is quite similar to free logic with negative semantics, known as NFL, and Scott's system, in that in all three systems an equational formula  $t = u$  implies existence of the objects denoted by  $t$  and  $u$ . Moreover, in our system as well as in NFL, all the functions are *strict* (in Scott's terminology), that is,  $Ef(a_1, \dots, a_n)$  implies  $Ea_j$  ( $j = 1, \dots, n$ ). The rules for universal and existential quantification are also similar.

On the other hand, our  $R\neg$  rule is more restrictive than that of either NFL or Scott's system. Also, we do not have *full preservation of equality* for functions,  $a_1 = b_1, \dots, a_n = b_n \rightarrow f(a_1, \dots, a_n) = f(b_1, \dots, b_n)$ , as axioms (although that can be proved as a theorem). Furthermore, we do not have *totality* of functions,  $Ea_1, \dots, Ea_n \rightarrow Ef(a_1, \dots, a_n)$ , as axioms (again, that can be proved as a theorem). The only functions for which we have totality as axioms are the binary successor functions  $s_0$  and  $s_1$ ; however, we can prove the totality of the other functions from the binary successor functions via the  $\Sigma_0^b$ -PIND rule.

As we will see in the next section, the restrictions we have imposed on our system are crucial to the provability in  $S_2$  of the consistency of strictly  $i$ -normal proofs in  $S_2^{-1}E$ .

### 3. $S_2^{i+2}$ PROOF OF CONSISTENCY OF STRICTLY $i$ -NORMAL PROOFS

In this section, we define *i-normal formula* and *strictly i-normal proof*, and in  $S_2^{i+2}$  we prove the consistency of strictly  $i$ -normal proofs in  $S_2^{-1}E$ . The consistency proof is based on the facts that we can produce a  $\Sigma_i^b$  formula that constitutes a truth definition for  $i$ -normal formulae and we can apply the  $\Sigma_{i+2}^b$ -PIND rule to prove the soundness of strictly  $i$ -normal proofs in  $S_2^{-1}E$ . The idea is that to use a term  $t$  in an  $S_2^{-1}E$  proof, we first need to prove that  $Et$  holds. To do this, we show that for a given assignment  $\rho$  of values to the variables in  $t$ , the value of  $t$  is bounded by the size of the proof of  $Et$  plus the size of  $\rho$ . Therefore, we can define a valuation function for terms and a truth definition for the formulae in the proof. Once we obtain the truth definition, consistency is easy to prove.

In Subsection 3.1, we describe how we assign Gödel numbers to formulae and proofs. In Subsection 3.2, we introduce a “bounded” valuation for terms and prove its basic properties. In Subsection 3.3, we introduce a “bounded” truth definition for quantifier-free formulae and then extend it to  $i$ -normal formulae. In Subsection 3.4, we introduce strictly  $i$ -normal proofs in  $S_2^{-1}E$ , and we prove the soundness of such proofs with respect to the truth definition given in Subsection 3.3.

In what follows,  $i$  is assumed to be a fixed integer from the set  $\{-1, 0, 1, 2, \dots\}$ .

**3.1. Gödel numbers for formulae and proofs.** Before proceeding to the main topics treated in this section, we need to explain how we assign Gödel numbers to the formulae of  $S_2^iE$  and to  $S_2^iE$  proofs.

Since our proofs are structured as proof trees (nested sequences of formulae and rules of inference), we first assign Gödel numbers to the symbols which appear in the formulae and rules of inference of  $S_2^iE$ , using a different natural number for each symbol. Then we assign Gödel numbers to finite sequences of natural numbers, and we apply this Gödel numbering to the formulae of  $S_2^iE$  and the nodes of proof trees.

In assigning Gödel numbers to sequences of natural numbers, we follow Buss’ method [5]. The Gödel number of the sequence  $\langle u_1, \dots, u_n \rangle$  is determined by the following procedure.

- (1) For every  $j$  ( $1 \leq j \leq n$ ), let  $u'_j$  be the natural number whose binary representation is obtained by inserting 0 between consecutive bits in the binary representation of  $u_j$ , and then appending 0 to the resulting bit string. (By convention,  $0' = 0$ .) For example, if the binary representation of  $u_j$  is 11, then that of  $u'_j$  is 1010; and if the binary representation of  $u_j$  is 101, then that of  $u'$  is 100010.
- (2) Define the Gödel number of  $\langle u_1, \dots, u_n \rangle$  to be  $u'_1 \oplus 3 \oplus u'_2 \oplus 3 \oplus \dots \oplus 3 \oplus u'_{n-1} \oplus 3 \oplus u'_n$ . Since the binary representation of the number 3 is 11, the Gödel number of the sequence  $\langle u_1, \dots, u_n \rangle$  is the natural number whose binary representation is obtained by inserting 11 between the binary representations of  $u'_j$  and  $u'_{j+1}$  ( $j = 1, \dots, n-1$ ).
- (3) The Gödel number of  $\langle u_1 \rangle$  is defined to be  $u'_1$ , and the Gödel number of the empty sequence  $\langle \rangle$  defined to be 0.

The notation  $\lceil \cdot \rceil$  is used for the Gödel number of any symbol or sequence. For example, the Gödel number of the formula  $A(a)$  is denoted by  $\lceil A(a) \rceil$ .

**3.2. Valuation of terms inside  $S_2^i$ .** In this subsection, we define a valuation function for the terms of  $S_2^i E$ . Our strategy in defining this valuation is to attach values to the nodes of a tree which is made up of all the subterms of the given term, and then to define the value of the term as the value attached to the root of the tree (the node that represents the entire term). We can view construction of this tree as a process of computation we undertake to obtain the value of its root.

**Definition 3.1.** Let  $\rho$  be a finite sequence of pairs  $([x], n)$  where  $x$  is a variable of  $S_2^i E$  and  $n \in \mathbb{N}$ .

- (1)  $\rho$  is an *environment* if, for every variable  $x$  of  $S_2^i E$ , there is at most one  $n \in \mathbb{N}$  with  $([x], n) \in \rho$ .
- (2) If  $\rho$  is an environment, we write  $\rho(x) = n$  to denote that  $([x], n) \in \rho$ , and we sometimes write  $\rho(\vec{a})$  to denote the sequence  $\rho(a_1), \dots, \rho(a_k)$  for a finite sequence  $\vec{a} \equiv a_1, \dots, a_k$  of variables of  $S_2^i E$ .
- (3) Let  $\sigma$  be a term, a formula, or a sequent.  $\rho$  is an *environment for  $\sigma$*  if  $\rho$  is an environment and, for every variable  $x$  that occurs free in  $\sigma$ , there is a pair  $([x], n)$  in  $\rho$  (for some  $n \in \mathbb{N}$ ).
- (4) If  $\rho$  is an environment, and if  $x$  is a variable and  $n \in \mathbb{N}$ , then  $\rho[x \mapsto n]$  denotes environment obtained from  $\rho$  by replacing the pair  $([x], \rho(x))$  with  $([x], n)$  if there is some  $m \in \mathbb{N}$  with  $([x], m)$  in  $\rho$ , and by adding the pair  $([x], n)$  to  $\rho$  otherwise.
- (5)  $\text{Env}$  denotes the ternary relation that holds of precisely the triples  $(\rho', [\sigma], u)$  where  $\sigma$  is a term, a formula, or a sequent;  $\rho'$  is an environment for  $\sigma$ ;  $u \in \mathbb{N}$ ; for every variable  $x$  of  $S_2^i E$ , there is a pair  $([x], n)$  in  $\rho'$  (for some  $n \in \mathbb{N}$ ) if and only if  $x$  occurs free in  $\sigma$ ; and  $\rho'(x) \leq u$  for every variable  $x$  that occurs free in  $\sigma$ . From this point on, we identify environments with their Gödel numbers; therefore, we regard  $\text{Env}$  as a ternary relation on  $\mathbb{N}$ .
- (6) Let  $u \in \mathbb{N}$ , and let  $\sigma$  be a term, a formula, or a sequent.  $\text{BdEnv}([\sigma], u)$  denotes the greatest  $m \in \mathbb{N}$  which is (the Gödel number of) an environment  $\rho'$  such that  $\text{Env}(\rho', [\sigma], u)$  holds.

Although an environment is a sequence—not a set—of pairs, from this point on we make no reference to the order of the pairs in an environment.

For a term, formula, or sequent  $\sigma$  and natural numbers  $\rho, u$ , the relation  $\text{Env}(\rho, [\sigma], u)$  and the function  $\text{BdEnv}([\sigma], u)$  are  $\Sigma_1^b$ -definable. Also, if  $t$  is a term and  $\rho$  is an environment for  $t$ , we can extend  $\rho$  to  $t$  by recursion on the construction of  $t$ , as follows:

$$\rho(0) := 0 \tag{3.1}$$

$$\rho(f(t_1, \dots, t_k)) := f(\rho(t_1), \dots, \rho(t_k)) \tag{3.2}$$

For a fixed term  $t$ , the function that, given an arbitrary environment  $\rho'$  for  $t$ , maps  $\rho'$  to the value of  $\rho'(t)$  obtained by this method is definable in  $S_2^1$ . For a fixed environment  $\rho$ , however, the function that, given an arbitrary term  $t'$  such that  $\rho$  is an environment for  $t'$ , maps  $t'$  to the value of  $\rho(t')$  obtained by this method is not definable in  $S_2$ , since there are sequences of such terms  $t'$  for which the values of  $\rho(t')$  increase exponentially in the length of  $t'$ .

**Definition 3.2.** Let  $t$  be a term of  $S_2^i E$ , let  $\rho$  be an environment for  $t$ , and let  $u \in \mathbb{N}$ . A  $\rho$ -*valuation tree for  $t$  which is bounded by  $u$*  is a tree  $w$  that satisfies the following conditions.

- (1) Every node of  $w$  is of the form  $\langle [t_j], c \rangle$  where  $t_j$  is a subterm of  $t$ ,  $c \in \mathbb{N}$ , and  $c \leq u$ .

- (2) Every leaf of  $w$  is either  $\langle [0], 0 \rangle$  or  $\langle [a], \rho(a) \rangle$  for some variable  $a$  in the domain of  $\rho$ .
- (3) The root of  $w$  is  $\langle [t], c \rangle$  for some  $c \leq u$ .
- (4) If  $\langle [f(t_1, \dots, t_n)], c \rangle$  is a node of  $w$ , then the children of this node are the nodes  $\langle [t_1], d_1 \rangle, \dots, \langle [t_n], d_n \rangle$  which satisfy the condition  $c = f(d_1, \dots, d_n)$ .

If the root of a  $\rho$ -valuation tree  $w$  for  $t$  is  $\langle [t], c \rangle$ , we say *the value of  $w$  is  $c$* .

The statement that  $t$  converges to the value  $c$  (and  $c \leq u$ ) is defined by the  $\Sigma_1^b$  formula which expresses that the following relation (which we denote by  $v([t], \rho) \downarrow_u c$ ) holds: “ $\exists w \leq s([t], u)$  such that  $w$  is (the Gödel number of) a  $\rho$ -valuation tree for  $t$  which is bounded by  $u$  and has root  $\langle [t], c \rangle$ ,” where  $s([t], u)$  is a term whose Gödel number bounds (the Gödel numbers of) all  $\rho$ -valuation trees for  $t$  which are bounded by  $u$ .

Here are some simple facts about the relation  $v([t], \rho) \downarrow_u c$ .

**Lemma 3.3.** *For a term  $t$ , an environment  $\rho$  for  $t$ , and  $u \in \mathbb{N}$ , the following statements are provable in  $S_2^1$ .*

- (1) *If  $v([t], \rho) \downarrow_u c$  and  $v([t], \rho) \downarrow_u c'$ , then  $c = c'$ .*
- (2) *If  $f$  is an  $n$ -ary function such that  $t \equiv ft_1 \dots t_n$  for terms  $t_1, \dots, t_n$  and we have  $v([f(t_1, \dots, t_n)], \rho) \downarrow_u c$ , then there exist  $d_1, \dots, d_n$  such that  $f(d_1, \dots, d_n) = c$  and  $v([t_1], \rho) \downarrow_u d_1, \dots, v([t_n], \rho) \downarrow_u d_n$ .*
- (3)  *$v([0], \rho) \downarrow_0 0$*
- (4) *For every variable  $a$  that occurs in  $t$ ,  $v([a], \rho) \downarrow_{\rho(a)} \rho(a)$ .*
- (5) *If  $v([t], \rho) \downarrow_u c$  and  $\rho' \subset \rho$  is an environment for  $t$ ,  $v([t], \rho') \downarrow_u c'$  and  $c = c'$ .*

Lemmata 3.4, 3.5 and 3.6 are key properties of  $v$  that are used in our proof of the consistency of strictly  $i$ -normal proofs (Subsection 3.4). Lemma 3.4 states that substitution is provably equivalent to assignment. Lemma 3.5 states that the relation  $v([t], \rho) \downarrow_u c$  is closed upward with respect to  $u$ . Lemma 3.6 states that if  $u$  is sufficiently large, the relation  $v([t], \rho) \downarrow_u c$  coincides with the valuation of  $t$ .

**Lemma 3.4.** *For terms  $t(\vec{a}, a)$  and  $t'(\vec{a})$  such that the variable  $a$  occurs in  $t$  and is not in  $\vec{a}$ , the following is provable in  $S_2^1$ .*

$$v([t(\vec{a}, t'(\vec{a}))], \rho) \downarrow_u c \leftrightarrow \exists c' \leq u, v([t'(\vec{a})], \rho) \downarrow_u c' \wedge v([t(\vec{a}, a)], \rho[a \mapsto c']) \downarrow_u c$$

*Proof.* By induction on the contraction of  $t$ . □

**Lemma 3.5.** *The following statement is provable in  $S_2^1$ : “If  $v([t], \rho) \downarrow_u c$  and  $u < u'$ , then  $v([t], \rho) \downarrow_{u'} c$ .”*

*Proof.* The lemma holds since if  $\rho$ -valuation tree  $w$  is bounded by  $u$ , then it is bounded by  $u'$ . □

**Lemma 3.6.** *Let  $t$  be a term of  $S_2^i E$ , let  $t_1, \dots, t_m$  be an enumeration of all the subterms of  $t$ . Then the following statement is provable in  $S_2^1$ : “For any environment  $\rho$  for  $t$  and  $u \in \mathbb{N}$ , if  $\rho(t_j) \leq u$  for every  $j$ , then  $v([t], \rho) \downarrow_u \rho(t)$  holds; and if  $\rho(t_j) > u$  for some  $j$ , then  $v([t], \rho) \downarrow_u c$  does not hold for any natural number  $c \leq u$ .”*

*Proof.* If  $\rho(t_j) \leq u$  for every  $j$ , we can construct a  $\rho$ -valuation tree for  $t$  which is bounded by  $u$ , by induction (outside of  $S_2^1$ ) on  $t$ . If  $\rho(t_j) > u$  for some  $j$ , then there is no  $\rho$ -valuation tree for  $t$  which is bounded by  $u$  and this fact can be proved in  $S_2^1$ . □



**3.3. Truth definition inside  $S_2^i$ .** In this subsection, we give a “bounded” truth definition for quantifier-free formulae, and then we extend the definition to the formulae  $i$ -normal formulae [16], [17], [18].

First, we present a truth definition for quantifier-free formulae. Since logical symbols can be arbitrarily nested, we follow the same strategy that was used in our definition of valuation for terms. We attach a truth value to each node of a subformula tree, and we define the value attached to the root (the node that represents the entire formula) as the truth value of the formula.

**Definition 3.7.** Let  $A$  be a quantifier-free formula of  $S_2^{-1}E$ , let  $\rho$  be an environment for  $A$ , and let  $u \in \mathbb{N}$ . A  $\rho$ -truth tree for  $A$  which is bounded by  $u$  is a tree  $w$  that satisfies the following conditions.

Every leaf of  $w$  has one of the following five forms (where in each form the possible values of  $\epsilon$  are 0 and 1):  $\langle \lceil t_1 \leq t_2 \rceil, \epsilon \rangle$ ,  $\langle \lceil t_1 \not\leq t_2 \rceil, \epsilon \rangle$ ,  $\langle \lceil t_1 = t_2 \rceil, \epsilon \rangle$ ,  $\langle \lceil t_1 \neq t_2 \rceil, \epsilon \rangle$ ,  $\langle \lceil Et \rceil, \epsilon \rangle$ .

For a leaf of the form  $\langle \lceil t_1 \leq t_2 \rceil, \epsilon \rangle$ ,  $\epsilon = 1$  if  $\exists c_1, c_2 \leq u$ ,  $v(\lceil t_1 \rceil, \rho) \downarrow_u c_1$ ,  $v(\lceil t_2 \rceil, \rho) \downarrow_u c_2$ , and  $c_1 \leq c_2$ ; otherwise,  $\epsilon = 0$ .

For a leaf of the form  $\langle \lceil t_1 \not\leq t_2 \rceil, \epsilon \rangle$ ,  $\epsilon = 1$  if  $\exists c_1, c_2 \leq u$ ,  $v(\lceil t_1 \rceil, \rho) \downarrow_u c_1$ ,  $v(\lceil t_2 \rceil, \rho) \downarrow_u c_2$ , and  $c_1 \not\leq c_2$ ; otherwise,  $\epsilon = 0$ .

The conditions that must be satisfied by a leaf of the form  $\langle \lceil t_1 = t_2 \rceil, \epsilon \rangle$  or  $\langle \lceil t_1 \neq t_2 \rceil, \epsilon \rangle$  are the obvious analogues of those for  $\langle \lceil t_1 \leq t_2 \rceil, \epsilon \rangle$  and  $\langle \lceil t_1 \not\leq t_2 \rceil, \epsilon \rangle$ , respectively.

For a leaf of the form  $\langle \lceil Et \rceil, \epsilon \rangle$ ,  $\epsilon = 1$  if  $\exists c \leq u$ ,  $v(\lceil t \rceil, \rho) \downarrow_u c$ ; otherwise,  $\epsilon = 0$ .

Every intermediate node  $r$  of  $w$  is of the form  $\langle \lceil A_1 \wedge A_2 \rceil, \epsilon \rangle$  or  $\langle \lceil A_1 \vee A_2 \rceil, \epsilon \rangle$ , where the children of  $r$  are the nodes  $\langle \lceil A_1 \rceil, \epsilon_1 \rangle$  and  $\langle \lceil A_2 \rceil, \epsilon_2 \rangle$ .

For a node of the form  $\langle \lceil A_1 \wedge A_2 \rceil, \epsilon \rangle$ ,  $\epsilon = 1$  if  $\epsilon_1 = 1$  and  $\epsilon_2 = 1$ ; otherwise,  $\epsilon = 0$ .

For a node of the form  $\langle \lceil A_1 \vee A_2 \rceil, \epsilon \rangle$ ,  $\epsilon = 1$  if  $\epsilon_1 = 1$  or  $\epsilon_2 = 1$ ; otherwise,  $\epsilon = 0$ .

The root of  $w$  is  $\langle \lceil A \rceil, \epsilon \rangle$  for some  $\epsilon \in \{0, 1\}$ .

The truth of a quantifier-free formula  $A$  is defined by the  $\Sigma_1^b$  formula  $T_{-1}(u, \lceil A \rceil, \rho)$  which expresses that “ $\exists w \leq s(\lceil A \rceil, u)$  such that  $w$  is (the Gödel number of) a  $\rho$ -truth tree for  $A$  which is bounded by  $u$  and has root  $\langle \lceil A \rceil, 1 \rangle$ ,” where  $s(\lceil A \rceil, u)$  is a term which bounds (the Gödel numbers of) all  $\rho$ -truth trees for  $A$  which are bounded by  $u$ .

We can prove several basic properties of the truth definition  $T_{-1}$ .

**Lemma 3.8.** *The following statements are provable in  $S_2^1$ .*

- (1)  $T_{-1}(u, \lceil t_1 \leq t_2 \rceil, \rho) \leftrightarrow \exists c_1, c_2 \leq u, v(\lceil t_1 \rceil, \rho) \downarrow_u c_1 \wedge v(\lceil t_2 \rceil, \rho) \downarrow_u c_2 \wedge c_1 \leq c_2$
- (2)  $T_{-1}(u, \lceil t_1 \not\leq t_2 \rceil, \rho) \leftrightarrow \exists c_1, c_2 \leq u, v(\lceil t_1 \rceil, \rho) \downarrow_u c_1 \wedge v(\lceil t_2 \rceil, \rho) \downarrow_u c_2 \wedge c_1 \not\leq c_2$
- (3)  $T_{-1}(u, \lceil t_1 = t_2 \rceil, \rho) \leftrightarrow \exists c_1, c_2 \leq u, v(\lceil t_1 \rceil, \rho) \downarrow_u c_1 \wedge v(\lceil t_2 \rceil, \rho) \downarrow_u c_2 \wedge c_1 = c_2$
- (4)  $T_{-1}(u, \lceil t_1 \neq t_2 \rceil, \rho) \leftrightarrow \exists c_1, c_2 \leq u, v(\lceil t_1 \rceil, \rho) \downarrow_u c_1 \wedge v(\lceil t_2 \rceil, \rho) \downarrow_u c_2 \wedge c_1 \neq c_2$
- (5)  $T_{-1}(u, \lceil Et \rceil, \rho) \leftrightarrow \exists c \leq u, v(\lceil t \rceil, \rho) \downarrow_u c$
- (6)  $T_{-1}(u, \lceil A_1 \wedge A_2 \rceil, \rho) \leftrightarrow T_{-1}(u, \lceil A_1 \rceil, \rho) \wedge T_{-1}(u, \lceil A_2 \rceil, \rho)$
- (7)  $T_{-1}(u, \lceil A_1 \vee A_2 \rceil, \rho) \leftrightarrow T_{-1}(u, \lceil A_1 \rceil, \rho) \vee T_{-1}(u, \lceil A_2 \rceil, \rho)$
- (8)  $T_{-1}(u, \lceil A(\vec{a}, t(\vec{a})) \rceil, \rho) \leftrightarrow \exists c \leq u, v(\lceil t(\vec{a}) \rceil, \rho) \downarrow_u c \wedge T_{-1}(u, \lceil A(\vec{a}, a) \rceil, \rho[a \mapsto c])$
- (9) *If  $\rho' \subset \rho$  is an environment for  $A$ ,  $T_{-1}(u, \lceil A \rceil, \rho') \leftrightarrow T_{-1}(u, \lceil A \rceil, \rho)$*

Lemmata 3.9, 3.10, 3.11, and 3.12 are key properties of  $T_{-1}$  that are used in our consistency proof. Lemma 3.9 states that substitution is provably equivalent to assignment. Lemma 3.10 states that  $T_{-1}$  is closed upward with respect to  $u$ . Lemma 3.11 states a reflection principle for  $T_{-1}$ . Lemma 3.12 states the law of the excluded middle.

**Lemma 3.9.** *Let  $A$  be a quantifier-free formula in which the variable  $a$  occurs, and let  $t$  be a term. Then the following statement is provable in  $S_2^1$ : “ $T_{-1}(u, [A(t)], \rho)$  if and only if there exists  $c \leq u$  such that  $v([t], \rho) \downarrow_u c$  and  $T_{-1}(u, [A(a)], \rho[a \mapsto c])$ .”*

**Lemma 3.10.** *For a quantifier-free formula  $A$ , it is provable in  $S_2^1$  that  $T_{-1}(u, [A], \rho), u < u' \rightarrow T_{-1}(u', [A], \rho)$ .*

*Proof.* Induction on the construction of  $A$ , using Lemma 3.5 for basic formulae  $A$ , and clauses 6 and 7 of Lemma 3.8 for other quantifier-free formulae.  $\square$

We can prove that  $T_{-1}$  is a truth definition by showing a kind of reflection principle for  $T_{-1}$ .

**Lemma 3.11.** *Let  $A(\vec{a})$  be a quantifier-free formula, let  $A'(\vec{a})$  be the interpretation of  $A(\vec{a})$  which is given in Definition 2.7 and let  $t_1, \dots, t_m$  be an enumeration of all the subterms of the terms that occur in  $A(\vec{a})$ . Then the following statement is provable in  $S_2^1$ : “If  $\rho$  be an environment for  $A$ ,  $u \in \mathbb{N}$  and  $\rho(t_j) \leq u$  for every  $j$ , then  $T_{-1}(u, [A(\vec{a})], \rho) \leftrightarrow A'(\rho(\vec{a}))$ ; and if  $\rho(t_j) > u$  for some  $j$ , then  $\neg T_{-1}(u, [A(\vec{a})], \rho)$ .”*

*Proof.* Induction on the construction of  $A$ , using Lemma 3.6 for basic formulae  $A$ .  $\square$

**Lemma 3.12.** *If  $p$  is  $=$  or  $\leq$ , it is provable in  $S_2^1$  that if  $T_{-1}(u, [Et_i], \rho)$  for  $i \in \{1, 2\}$  hold, then either  $T_{-1}(u, [p(t_1, t_2)], \rho)$  or  $T_{-1}(u, [\neg p(t_1, t_2)], \rho)$  holds.*

Next, we would like to present a truth definition for  $\Sigma_i^b$  formulae. However, since it is technically difficult to do this for general  $\Sigma_i^b$  formulae, we restrict our definition to  $i$ -normal formulae. Since  $i \in \{-1, 0, 1, 2, \dots\}$ , we have  $-1$ -normal formulae,  $0$ -normal formulae,  $1$ -normal formulae,  $2$ -normal formulae, and so on.

**Definition 3.13.** Let  $i \geq -1$ , and let  $A(\vec{a})$  be a formula.

If  $i = -1$ ,  $A(\vec{a})$  is *pure  $-1$ -normal* if  $A(\vec{a})$  is quantifier free.

If  $i \geq 0$ ,  $A(\vec{a})$  is *pure  $i$ -normal* if it is of the form

$$\exists x_1 \leq t_1(\vec{a}) \forall x_2 \leq t_2(\vec{a}, x_1) \cdots$$

$$Q_i x_i \leq t_i(\vec{a}, x_1, \dots, x_{i-1}) Q_{i+1} x_{i+1} \leq |t_{i+1}(\vec{a}, x_1, \dots, x_i)|. A_0(\vec{a}, x_1, \dots, x_{i+1}),$$

where  $Q_i$  is  $\forall$  if  $i$  is even, and  $\exists$  if  $i$  is odd (and vice versa for  $Q_{i+1}$ ), and  $A_0(\vec{a}, x_1, \dots, x_{i+1})$  is quantifier free and does not contain the predicate  $E$ . Note that if  $i \geq 0$ , then a pure  $i$ -normal formula has at least one quantifier, so quantifier-free formulae are not pure  $i$ -normal. Also note that unlike the variables  $x_1, \dots, x_i$  (which are bounded by the values of the terms  $t_1, \dots, t_i$ , respectively), the value of the variable  $x_{i+1}$  is bounded by  $|t_{i+1}(\vec{a}, x_1, \dots, x_i)|$ , that is, by the number of bits in the binary representation of the value of the term  $t_{i+1}(\vec{a}, x_1, \dots, x_i)$ .

If  $i = -1$ ,  $A(\vec{a})$  is  *$i$ -normal* if it is quantifier free.

If  $i \geq 0$ ,  $A(\vec{a})$  is  *$i$ -normal* if it is a subformula of a pure  $i$ -normal formula or is  $Et$  for some term  $t$ . In other words,  $A(\vec{a})$  is either an  $E$ -form, a quantifier-free formula that does not contain  $E$ , or a formula of the form

$$Q_j x_j \leq t_j(\vec{a}, x_1, \dots, x_{j-1}) \cdots Q_i x_i \leq t_i(\vec{a}, x_1, \dots, x_{i-1}) \\ Q_{i+1} x_{i+1} \leq |t_{i+1}(\vec{a}, x_1, \dots, x_i)|. A_0(\vec{a}, x_1, \dots, x_{i+1}), \quad (3.3)$$

where  $A_0(\vec{a}, x_1, \dots, x_{i+1})$  is quantifier free and does not contain  $E$ ;  $1 \leq j \leq i+1$ ; and for every  $k$  with  $j \leq k \leq i+1$ ,  $Q_k$  is either  $\forall$  or  $\exists$ , according as  $k$  is even or odd. If  $j = i+1$ , the above formula is  $Q_{i+1}x_{i+1} \leq |t_{i+1}(\vec{a}, x_1, \dots, x_i)|.A_0(\vec{a}, x_1, \dots, x_{i+1})$ .

The following is a truth definition  $T_i(u, [B], \rho)$  for  $i$ -normal formulae  $B$ . First, we define a truth definition  $T_{i,l}$  for  $i$ -normal forms with  $l$  quantifiers.

**Definition 3.14.** Let  $i \geq -1$ , let  $B$  be an  $i$ -normal formula with  $l$  quantifiers. Note that  $0 \leq l \leq i+1$ . We define  $T_{i,l}(u, [B], \rho)$  by recursion on  $l$  in the meta-language.

If  $l = 0$ , then  $B$  is quantifier free, so  $T_i(u, [B], \rho) \equiv T_{-1}(u, [B], \rho)$ .

If  $l \geq 1$ , then

$$B \equiv Q_j x_j \leq t.A(\vec{a}, x_1, \dots, x_j),$$

where  $j = i+2-l$ ;  $t \equiv t_j(\vec{a}, x_1, \dots, x_{j-1})$  if  $j < i+1$ , and  $t \equiv |t_{i+1}(\vec{a}, x_1, \dots, x_i)|$  if  $j = i+1$ ; and  $A(\vec{a}, x_1, \dots, x_j)$  is an  $i$ -normal formula with  $l-1$  quantifiers. Assume that we have defined  $T_{i,l-1}(u, [C], \rho)$  for all  $i$ -normal formulae  $C$  with  $l-1$  quantifiers. We define  $T_{i,l}(u, [B], \rho)$  to be the following formula.

$$\exists c \leq u, v([t], \rho) \downarrow_u c \wedge Q_j d_j \leq c.T_i(u, [A(\vec{a}, x_1, \dots, x_j)], \rho[x_j \mapsto d_j])$$

Then, let  $\text{INQ}([B], l)$  be a formula which represents “ $B$  is an  $i$ -normal form with  $l$  quantifiers”. we define  $T_i(u, [B], \rho)$  as

$$\{\text{INQ}([B], 0) \supset T_{i,0}(u, [B], \rho)\} \vee \dots \vee \{\text{INQ}([B], i+1) \supset T_{i,i+1}(u, [B], \rho)\} \quad (3.4)$$

Since we can contract successive  $\exists$  quantifiers into a single  $\exists$  quantifier,  $T_i(u, [B], \rho)$  is  $\Sigma_{i+1}^b$ .

Lemmata 3.15 and 3.16 are key properties of  $T_i$  that are used in our consistency proof. Lemma 3.15 states that substitution is provably equivalent to assignment. Lemma 3.16 states that  $T_i$  is closed upward with respect to  $u$ .

**Lemma 3.15.** *Let  $A$  be an  $i$ -normal formula in which the variable  $a$  occurs, and let  $t$  be a term. Then the following statement is provable in  $S_2^1$ : “ $T_i(u, [A(t)], \rho)$  if and only if there exists  $c \leq u$  such that  $v([t], \rho) \downarrow_u c$  and  $T_i(u, [A(a)], \rho[a \mapsto c])$ .”*

*Proof.* Easy consequence of Lemma 3.9. □

Similarly, we can use Lemma 3.10 to prove upward closedness of  $T_i$  with respect to  $u$ .

**Lemma 3.16.** *For an  $i$ -normal formula  $A$ , it is provable in  $S_2^1$  that  $T_i(u, [A], \rho) \wedge u < u' \rightarrow T_i(u', [A], \rho)$ .*

As the clause 9 of Lemma 3.8, if  $\rho$  contains variables other than free variables which occurs in  $A$ , we can ignore such variables.

**Lemma 3.17.** *For an  $i$ -normal formula  $A$ , environments  $\rho$  for  $A$  and  $\rho \subset \rho'$ , it is provable in  $S_2^1$  that  $T_i(u, [A], \rho) \leftrightarrow T_i(u, [A], \rho')$ .*

By definition of  $T_i$ , we see that we can take the outermost quantifier of an  $i$ -normal formula whose Gödel number is the second argument of  $T_i$  and move it to the outside of  $T_i$ .

**Lemma 3.18.** *Let  $\forall x \leq t.A(x)$  and  $\exists x \leq t.A(x)$  be  $i$ -normal formulae with at least one quantifier (and whose outermost quantifier is of the indicated type). Then the following is*

provable in  $S_2^1$ .

$$T_i(u, [\forall x \leq t. A(x)], \rho) \Leftrightarrow \exists c \leq u, v([\![t]\!], \rho) \downarrow_u c \wedge \forall d \leq c. T_i(u, [A(a)], \rho[a \mapsto d]) \quad (3.5)$$

$$T_i(u, [\exists x \leq t. A(x)], \rho) \Leftrightarrow \exists c \leq u, v([\![t]\!], \rho) \downarrow_u c \wedge \exists d \leq c. T_i(u, [A(a)], \rho[a \mapsto d]) \quad (3.6)$$

**3.4. Strictly  $i$ -normal proofs and their consistency.** Since we have defined truth for  $i$ -normal formulae only, we can define soundness for only those proofs that consist entirely of  $i$ -normal formulae. We call such proofs *strictly  $i$ -normal*, and we use the term *strictly  $i$ -normal proof tree for  $\Gamma \rightarrow \Delta$*  to mean a tree  $w$  which represents a proof of  $\Gamma \rightarrow \Delta$ . In such a tree, every node  $r$  has the form  $\langle R, [\Gamma_r \rightarrow \Delta_r], w_1, \dots, w_{l(R)} \rangle$ , where  $R$  is a name of inference,  $\Gamma_r \rightarrow \Delta_r$  is the conclusion of the inference, and  $l(R)$  is a number of premises of  $R$ . If  $l(R) = 0$ , the node has the form  $\langle R, [\Gamma_r \rightarrow \Delta_r] \rangle$ .

**Definition 3.19.** An  $S_2^{-1}E$  proof is *strictly  $i$ -normal* if all formulae contained in the proof are  $i$ -normal. The property “ $w$  is (the Gödel number of) a strictly  $i$ -normal proof tree for  $\Gamma \rightarrow \Delta$ ” is  $\Delta_1^b$ -definable. We write  $i\text{-Prf}(w, [\Gamma \rightarrow \Delta])$  for the  $\Delta_1^b$  formula that defines this property.

**Proposition 3.20.** Let  $\Gamma \rightarrow \Delta$  be a sequent comprised entirely of  $i$ -normal formulae, and let  $u, w \in \mathbb{N}$  such that  $i\text{-Prf}(w, [\Gamma \rightarrow \Delta])$  holds,  $w \leq u$ , and the binary representation of  $u$  is of the form  $11 \cdots 1$ , that is, all the bits are 1. Then for every node  $r$  of  $w$ , the following holds (where  $\rho$  denotes an environment as well as its Gödel number and  $\Gamma_r \rightarrow \Delta_r$  denotes the conclusion of the subproof which corresponds a node  $r$ ).

$$\begin{aligned} \forall \rho \leq \text{BdEnv}([\Gamma_r \rightarrow \Delta_r], u) & \left[ \text{Env}(\rho, [\Gamma_r \rightarrow \Delta_r], u) \supset \right. \\ & \left. \forall u' \leq u \ominus r \{ [\forall A \in \Gamma_r, T_i(u', [A], \rho)] \supset [\exists B \in \Delta_r, T_i(u' \oplus r, [B], \rho)] \} \right] \quad (3.7) \end{aligned}$$

Furthermore, this is derivable in  $S_2^{i+2}$ .

*Proof.* We prove the proposition by tree induction on  $r$ . Since the formula in (3.7) is  $\Pi_{i+2}^b$ , our proof can be carried out in  $S_2^{i+2}$ .

Let  $\rho \leq \text{BdEnv}([\Gamma_r \rightarrow \Delta_r], u)$ , assume that  $\text{Env}(\rho, [\Gamma_r \rightarrow \Delta_r], u)$ , and let  $u' \leq u \ominus r$ . Note that  $u \geq r$ , since  $u \geq w \geq r$ . Therefore,  $u' \oplus r \leq u \ominus r \oplus r \leq u$ , since all the bits in the binary representation of  $u$  are 1. We use this fact throughout this proof.

We prove that if  $\forall A \in \Gamma_r, T_i(u', [A], \rho)$  holds, then  $\exists B \in \Delta_r, T_i(u' \oplus r, [B], \rho)$ , by considering all possible forms for the last inference in the derivation of  $\Gamma_r \rightarrow \Delta_r$ .

Identity rule:

$$\overline{A \rightarrow A} \text{ Id} \quad (3.8)$$

Assume that  $T_i(u', [A], \rho)$ . Then, by Lemma 3.16,  $T_i(u' \oplus r, [A], \rho)$ . Hence,  $r$  satisfies (3.7).

Axiom rule:

$$\frac{}{\Gamma(\vec{s}(\vec{a})) \rightarrow \Delta(\vec{s}(\vec{a}))} \text{Ax}, \quad (3.9)$$

where  $\Gamma(\vec{s}(\vec{a})) \rightarrow \Delta(\vec{s}(\vec{a}))$  is a substitution instance of an axiom.

Since there are only finite many axioms, we use case analysis on the axiom which derives this substitution instance. Assume that  $\forall A \in \Gamma, T_i(u', \lceil A(\vec{s}(\vec{a})) \rceil, \rho)$ . Let  $\Gamma(\vec{b}) \rightarrow \Delta(\vec{b})$  be the axiom into which the substitution was made. By the assumption on  $\mathcal{A}$ , this axiom satisfies the boundedness conditions (Definition 2.9). Moreover, its standard interpretation (given in Definition 2.7) is derivable in  $S_2^{i+2}$ .

By the first boundedness condition, all the formulae in  $\Gamma$  and  $\Delta$  are basic. Let  $\vec{b} = b_1, \dots, b_l$  and  $\vec{s}(\vec{a}) = s_1(\vec{a}), \dots, s_l(\vec{a})$ , where  $s_k(\vec{a})$  is the term that was substituted for the variable  $b_k$  in the application of the axiom rule ( $k = 1, \dots, l$ ). By the second boundedness condition,  $b_k$  occurs in  $\Gamma$ , so by Lemma 3.15  $\exists d_k \leq u'$  such that  $v(\lceil s_k(\vec{a}) \rceil, \rho) \downarrow_{u'} d_k$  ( $k = 1, \dots, l$ ), hence  $\forall A \in \Gamma, T_i(u', \lceil A(\vec{b}) \rceil, \rho[\vec{b} \mapsto \vec{d}])$ .

Let  $\vec{t}_\Gamma(\vec{b})$  be the subterms of the terms that occur in  $\Gamma(\vec{b})$ , and let  $\vec{t}_\Delta(\vec{b})$  be the subterms of the terms that occur in  $\Delta(\vec{b})$ . Since all formulae occur in  $\Gamma$  and  $\Delta$  are basic,  $\vec{b}$  are all variables contained  $\vec{t}_\Gamma(\vec{b})$  and  $\Gamma(\vec{b})$ . Since the function symbol of  $S_2^i E$  is definable in  $S_2^1$ , we can view the terms in  $\vec{t}_\Gamma(\vec{b})$  and  $\vec{t}_\Delta(\vec{b})$  as terms of  $S_2^1$ . By the third boundedness condition, the relation

$$\max\{\vec{t}_\Delta(\vec{b})\} \leq \alpha \cdot \max\{\vec{t}_\Gamma(\vec{b})\} \quad (3.10)$$

is provable in  $S_2^1$ .

Since  $\forall A \in \Gamma, T_i(u', \lceil A(\vec{b}) \rceil, \rho[\vec{b} \mapsto \vec{d}])$ , we have  $\max\{\vec{t}_\Gamma(\vec{d})\} \leq u'$ . By Lemma 3.11, we have that, for every  $A$  in  $\Gamma$ ,  $A(\vec{d})$  is true (in the meta-language). Since  $\Gamma(\vec{d}) \rightarrow \Delta(\vec{d})$  holds (in the meta-language), there is some  $B$  in  $\Delta$  such that  $B(\vec{d})$  is true (in the meta-language). Since we can take  $\alpha$  to be 4, we have  $\max\{\vec{t}_\Delta(\vec{d})\} \leq 4 \cdot u' \leq u' \oplus r$ . Let  $\vec{c} = FV(B(\vec{b}))$ . Then  $T_i(u' \oplus r, \lceil B(\vec{c}) \rceil, \rho[\vec{c} \mapsto \vec{d}])$  holds by Lemma 3.18. By Lemma 3.5 and the fact that  $v(\lceil s_k(\vec{a}) \rceil, \rho) \downarrow_{u'} d_k$  ( $k = 1, \dots, l$ ), we obtain  $v(\lceil s_k(\vec{a}) \rceil, \rho) \downarrow_{u' \oplus r} d_k$ . Using that result and Lemma 3.15, we have  $T_i(u' \oplus r, \lceil B(\vec{s}(\vec{a})) \rceil, \rho)$ , so we are done.

Structural rules:

$$\frac{\begin{array}{c} \vdots \\ r_1 \\ \Gamma \rightarrow \Delta \end{array}}{A, \Gamma \rightarrow \Delta} \quad (3.11)$$

Assume that  $T_i(u', \lceil A \rceil, \rho)$  and  $\forall B \in \Gamma, T_i(u', \lceil B \rceil, \rho)$ . Let  $\rho_1$  be the subsequence of  $\rho$  such that  $\text{Env}(\rho_1, \lceil \Gamma_{r_1} \rightarrow \Delta_{r_1} \rceil, u)$ . Then  $\rho_1 \leq \text{BdEnv}(\lceil \Gamma_{r_1} \rightarrow \Delta_{r_1} \rceil, u)$  and by Lemma 3.17 we have  $\forall B \in \Gamma, T_i(u', \lceil B \rceil, \rho_1)$ . By the induction hypothesis applied to  $r_1$  together with the fact that  $u' \leq u \oplus r \leq u \oplus r_1$ , there is some  $B$  in  $\Delta$  such that  $T_i(u' \oplus r_1, \lceil B \rceil, \rho_1)$  holds, hence  $T_i(u' \oplus r_1, \lceil B \rceil, \rho)$  holds as well because of Lemma 3.17. By Lemma 3.16 and the fact that  $u' \oplus r_1 \leq u' \oplus r$ , we are done.

The proof for the other weakening rule is similar, and the proofs for the remaining structural rules are trivial.

$\neg$ -rules:

$$\frac{\begin{array}{c} \vdots \\ r_1 \\ \Gamma \rightarrow \Delta, p(t_1, t_2) \end{array}}{\neg p(t_1, t_2), \Gamma \rightarrow \Delta} \text{L} \neg, \quad (3.12)$$

where  $p$  is  $=$  or  $\leq$ .

Assume that  $T_i(u', [\neg p(t_1, t_2)], \rho)$  and  $\forall A \in \Gamma, T_i(u', [A], \rho)$ . Note that  $\text{Env}(\rho, [\Gamma_{r_1} \rightarrow \Delta_{r_1}], u)$  holds, because the variables that occur free in  $\Gamma_{r_1} \rightarrow \Delta_{r_1}$  are precisely those that occur free in  $\Gamma_r \rightarrow \Delta_r$ ; hence  $\rho \leq \text{BdEnv}([\Gamma_{r_1} \rightarrow \Delta_{r_1}], u)$ . Moreover,  $u' \leq u \ominus r \leq u \ominus r_1$ , so by the induction hypothesis applied to  $r_1$ , either  $T_i(u' \oplus r_1, [p(t_1, t_2)], \rho)$  holds or  $\exists B \in \Delta, T_i(u' \oplus r_1, [B], \rho)$ . By Lemma 3.16 and our assumption about  $\neg p(t_1, t_2)$ ,  $T_i(u' \oplus r_1, [\neg p(t_1, t_2)], \rho)$  holds, so  $T_i(u' \oplus r_1, [p(t_1, t_2)], \rho)$  cannot also hold. Hence  $\exists B \in \Delta, T_i(u' \oplus r_1, [B], \rho)$ . Since  $u' \oplus r_1 \leq u' \oplus r$ , by Lemma 3.16 we are done.

$$\frac{\begin{array}{c} \vdots \\ r_1 \\ p(t_1, t_2), \Gamma \rightarrow \Delta \end{array}}{Et_1, Et_2, \Gamma \rightarrow \Delta, \neg p(t_1, t_2)} \text{R} \neg, \quad (3.13)$$

where  $p$  is  $=$  or  $\leq$ .

Assume that  $T_i(u', [Et_1], \rho)$ ,  $T_i(u', [Et_2], \rho)$ , and  $\forall A \in \Gamma, T_i(u', [A], \rho)$ . If  $\exists B \in \Delta, T_i(u' \oplus r_1, [B], \rho)$ , we are done, so assume otherwise, that is, for every  $B \in \Delta$ ,  $T_i(u' \oplus r_1, [B], \rho)$  does not hold. Note that, just as in the proof for the  $\text{L} \neg$ -rule,  $\text{Env}(\rho, [\Gamma_{r_1} \rightarrow \Delta_{r_1}], u)$  holds and  $\rho \leq \text{BdEnv}([\Gamma_{r_1} \rightarrow \Delta_{r_1}], u)$ . Moreover,  $u' \leq u \ominus r \leq u \ominus r_1$ . Hence, by the induction hypothesis applied to  $r_1$ ,  $T_i(u', [p(t_1, t_2)], \rho)$  does not hold. By Lemma 3.12 and our assumption about  $Et_1$  and  $Et_2$ ,  $T_i(u', [\neg p(t_1, t_2)], \rho)$  does hold. Hence by Lemma 3.16 we are done.

$\wedge$ -rules:

$$\frac{\begin{array}{c} \vdots \\ r_1 \\ A, \Gamma \rightarrow \Delta \end{array}}{A \wedge B, \Gamma \rightarrow \Delta} \text{L} \wedge_1 \quad (3.14)$$

Assume that  $T_i(u', [A \wedge B], \rho)$  and  $\forall C \in \Gamma, T_i(u', [C], \rho)$ . Note that since  $A \wedge B$  is an  $i$ -normal formula, by Definition 3.13 it is quantifier free, hence  $T_i(u', [A \wedge B], \rho) \leftrightarrow T_{-1}(u', [A \wedge B], \rho)$ . By Lemma 3.8, we have  $T_i(u', [A], \rho)$ . Let  $\rho_1$  be the subsequence of  $\rho$  such that  $\text{Env}(\rho_1, [\Gamma_{r_1} \rightarrow \Delta_{r_1}], u)$ . Then we have  $T_i(u', [A], \rho_1)$  and  $\forall C \in \Gamma, T_i(u', [C], \rho_1)$  by Lemma 3.17; in addition,  $\rho_1 \leq \text{BdEnv}([\Gamma_{r_1} \rightarrow \Delta_{r_1}], u)$ . By the induction hypothesis applied to  $r_1$  together with the fact that  $u' \leq u \ominus r \leq u \ominus r_1$ , there is some  $D \in \Delta$  such that  $T_i(u' \oplus r_1, [D], \rho_1)$  holds. Then  $T_i(u' \oplus r_1, [D], \rho)$  holds as well by Lemma 3.17. Since  $u' \oplus r_1 \leq u' \oplus r$ , by Lemma 3.16 we are done.

The proof for the  $\text{L} \wedge_2$ -rule is similar.

$$\frac{\begin{array}{c} \vdots \\ r_1 \\ \Gamma \rightarrow \Delta, A \end{array} \quad \begin{array}{c} \vdots \\ r_2 \\ \Gamma \rightarrow \Delta, B \end{array}}{\Gamma \rightarrow \Delta, A \wedge B} \text{R} \wedge \quad (3.15)$$

Assume that  $\forall C \in \Gamma, T_i(u', [C], \rho)$ . Let  $\rho_1$  be the subsequence of  $\rho$  such that  $\text{Env}(\rho_1, [\Gamma_{r_1} \rightarrow \Delta_{r_1}], u)$ . Then we have  $\forall C \in \Gamma, T_i(u', [C], \rho_1)$ , and  $\rho_1 \leq \text{BdEnv}([\Gamma_{r_1} \rightarrow \Delta_{r_1}], u)$ . By the induction hypothesis applied to  $r_1$  together with the fact that  $u' \leq u \ominus r \leq u \ominus r_1$ , either  $\exists D \in \Delta, T_i(u' \oplus r_1, [D], \rho_1)$  or  $T_i(u' \oplus r_1, [A], \rho_1)$ .

Similarly, let  $\rho_2$  be the subsequence of  $\rho$  such that  $\text{Env}(\rho_2, [\Gamma_{r_2} \rightarrow \Delta_{r_2}], u)$ . Then we have  $\forall C \in \Gamma, T_i(u', [C], \rho_2)$ , and  $\rho_2 \leq \text{BdEnv}([\Gamma_{r_2} \rightarrow \Delta_{r_2}], u)$ . By the induction hypothesis applied to  $r_2$  together with the fact that  $u' \leq u \ominus r \leq u \ominus r_2$ , either  $\exists D \in \Delta, T_i(u' \oplus r_2, [D], \rho_2)$  or  $T_i(u' \oplus r_2, [B], \rho_2)$ .

If there exist  $j \in \{1, 2\}$  and  $D \in \Delta$  such that  $T_i(u' \oplus r_j, [D], \rho_j)$ , then  $T_i(u' \oplus r_j, [D], \rho)$  holds by Lemma 3.17. Thus we are done by Lemma 3.16 and the fact that  $u' \oplus r_j \leq u' \oplus r$ , so assume otherwise.

Then both  $T_i(u' \oplus r_1, [A], \rho_1)$  and  $T_i(u' \oplus r_2, [B], \rho_2)$  hold. Thus we have both  $T_i(u' \oplus r_1, [A], \rho)$  and  $T_i(u' \oplus r_2, [B], \rho)$ , by Lemma 3.17. Since  $u' \oplus r_1, u' \oplus r_2 \leq u' \oplus r$ , both  $T_i(u' \oplus r, [A], \rho)$  and  $T_i(u' \oplus r, [B], \rho)$  hold by Lemma 3.16. As noted in the proof for the  $L \wedge_1$ -rule, the formula  $A \wedge B$  is quantifier free. Thus by Lemma 3.8 and the definition of  $T_i$ , we have  $T_i(u' \oplus r, [A \wedge B], \rho)$ .

$\vee$ -rules: The proofs for  $\vee$ -rules are similar to the proofs for  $\wedge$ -rules.

Bounded  $\forall$ -rules:

$$\frac{\begin{array}{c} \vdots r_1 \\ A(t), \Gamma \rightarrow \Delta \end{array}}{t \leq s, \forall x \leq s. A(x), \Gamma \rightarrow \Delta} \text{Lb} \forall, \quad (3.16)$$

where the variable  $x$  does not occur in the term  $s$ .

Assume that  $T_i(u', [t \leq s], \rho)$ ,  $T_i(u', [\forall x \leq s. A(x)], \rho)$ , and  $\forall B \in \Gamma, T_i(u', [B], \rho)$ . Since  $T_i(u', [t \leq s], \rho)$  holds, there are  $c_0, d$  such that  $v([t], \rho) \downarrow_{u'} c_0$ ,  $v([s], \rho) \downarrow_{u'} d$ , and  $c_0 \leq d$ . By Lemma 3.18 and the fact that  $v([s], \rho) \downarrow_{u'} d$ , we have  $\forall c \leq d, T_i(u', [A(a)], \rho[a \mapsto c])$ . In particular,  $T_i(u', [A(a)], \rho[a \mapsto c_0])$ , since  $c_0 \leq d$ . If  $a$  occurs free in  $A(a)$ , we can apply Lemma 3.15 to  $A(a)$  and obtain  $T_i(u', [A(t)], \rho)$ . The conclusion is obvious if  $a$  does not occur free in  $A$ .

Note that  $\rho$  is an environment for  $\Gamma_{r_1} \rightarrow \Delta_{r_1}$ , since every variable that occurs free in  $\Gamma_{r_1} \rightarrow \Delta_{r_1}$  also occurs free in  $\Gamma_r \rightarrow \Delta_r$ , so let  $\rho_1$  be the subsequence of  $\rho$  such that  $\text{Env}(\rho_1, [\Gamma_{r_1} \rightarrow \Delta_{r_1}], u)$ . Then we have  $T_i(u', [A(t)], \rho_1)$  and  $\forall B \in \Gamma, T_i(u', [B], \rho_1)$ ; in addition,  $\rho_1 \leq \text{BdEnv}([\Gamma_{r_1} \rightarrow \Delta_{r_1}], u)$ . Thus there is some  $C$  in  $\Delta$  such that  $T_i(u' \oplus r_1, [C], \rho_1)$  holds, by the induction hypothesis applied to  $r_1$  together with the fact that  $u' \leq u \ominus r \leq u \ominus r_1$ . Hence we have  $T_i(u' \oplus r_1, [C], \rho)$ , because  $\rho_1$  is a subsequence of  $\rho$ . Since  $u' \oplus r_1 \leq u' \oplus r$ , by Lemma 3.16 we are done.

$$\frac{\begin{array}{c} \vdots r_1 \\ a \leq t, \Gamma \rightarrow \Delta, A(a) \end{array}}{Et, \Gamma \rightarrow \Delta, \forall x \leq t. A(x)} \text{Rb} \forall, \quad (3.17)$$

where neither the variable  $a$  nor the variable  $x$  occurs in the term  $t$ , and  $a$  does not occur free in  $\Gamma \rightarrow \Delta$ .

Assume that  $T_i(u', [Et], \rho)$  and  $\forall B \in \Gamma, T_i(u', [B], \rho)$ . Since  $T_i(u', [Et], \rho)$  holds, there is some  $c \leq u'$  such that  $v([t], \rho) \downarrow_{u'} c$ . Let  $d$  be any natural number such that  $d \leq c$ . Then  $T_i(u', [a \leq t], \rho[a \mapsto d])$  holds by Lemma 3.11 and the fact that  $T_i(u', [a \leq t], \rho[a \mapsto d]) \leftrightarrow T_{-1}(u', [a \leq t], \rho[a \mapsto d])$ . Furthermore, since  $a$  does not occur free in  $\Gamma$ , we have  $\forall B \in \Gamma, T_i(u', [B], \rho[a \mapsto d])$ .

Note that  $\rho[a \mapsto d]$  is an environment for  $\Gamma_{r_1} \rightarrow \Delta_{r_1}$ , since every variable other than  $a$  that occurs free in  $\Gamma_{r_1} \rightarrow \Delta_{r_1}$  also occurs free in  $\Gamma_r \rightarrow \Delta_r$ . Moreover,  $\rho[a \mapsto d](y) \leq u$  for every variable  $y$  that occurs free in  $\Gamma_{r_1} \rightarrow \Delta_{r_1}$ , since  $d \leq c \leq u' \leq u$ . Since  $\text{Env}(\rho, [\Gamma_r \rightarrow \Delta_r], u)$

$\Delta_r], u)$ ,  $\text{Env}(\rho[a \mapsto d], [\Gamma_{r_1} \rightarrow \Delta_{r_1}], u)$  holds. By the induction hypothesis applied to  $r_1$  together with the fact that  $u' \leq u \oplus r \leq u \oplus r_1$ , either  $\exists C \in \Delta, T_i(u' \oplus r_1, [C], \rho[a \mapsto d])$  or  $T_i(u' \oplus r_1, [A(a)], \rho[a \mapsto d])$ .

If there is some  $C$  in  $\Delta$  such that  $T_i(u' \oplus r_1, [C], \rho[a \mapsto d])$ , then we have  $T_i(u' \oplus r_1, [C], \rho)$ , since  $a$  does not occur free in  $D$ . Hence we are done by Lemma 3.16 and the fact that  $u' \oplus r_1 \leq u' \oplus r$ , so assume otherwise.

Then  $T_i(u' \oplus r_1, [A(a)], \rho[a \mapsto d])$  holds. Since  $d$  was an arbitrary natural number less than or equal to  $c$ , by Lemma 3.18 we have  $T_i(u' \oplus r_1, [\forall x \leq t. A(x)], \rho)$ . Since  $u' \oplus r_1 \leq u' \oplus r$ , we are done by Lemma 3.16.

Bounded  $\exists$ -rules:

$$\frac{\begin{array}{c} \vdots r_1 \\ a \leq t, A(a), \Gamma \rightarrow \Delta \end{array}}{\exists x \leq t. A(x), \Gamma \rightarrow \Delta} \text{Lb } \exists, \quad (3.18)$$

where neither the variable  $a$  nor the variable  $x$  occurs in the term  $t$ , and  $a$  does not occur free in  $\Gamma \rightarrow \Delta$ .

Assume that  $T_i(u', [\exists x \leq t. A(x)], \rho)$  and  $\forall B \in \Gamma, T_i(u', [B], \rho)$ . By Lemma 3.18, there exist  $c, d$  such that  $d \leq c$ ,  $v([t], \rho) \downarrow_{u'} c$ , and  $T_i(u', [A(a)], \rho[a \mapsto d])$ . Then  $T_i(u', [a \leq t], \rho[a \mapsto d])$  holds by Lemma 3.11 and the fact that  $T_i(u', [a \leq t], \rho[a \mapsto d]) \equiv T_{-1}(u', [a \leq t], \rho[a \mapsto d])$ . Since  $a$  does not occur free in  $\Gamma$ , we have  $\forall B \in \Gamma, T_i(u', [B], \rho[a \mapsto d])$ .

Note that  $\rho[a \mapsto d]$  is an environment for  $\Gamma_{r_1} \rightarrow \Delta_{r_1}$ , and that  $\rho[a \mapsto d](y) \leq u$  for every variable  $y$  that occurs free in  $\Gamma_{r_1} \rightarrow \Delta_{r_1}$ , since  $d \leq c \leq u' \leq u$ . Since  $\text{Env}(\rho, [\Gamma_r \rightarrow \Delta_r], u)$ ,  $\text{Env}(\rho[a \mapsto d], [\Gamma_{r_1} \rightarrow \Delta_{r_1}], u)$ . Thus by the induction hypothesis applied to  $r_1$  together with the fact that  $u' \leq u \oplus r \leq u \oplus r_1$ , there is some  $C$  in  $\Delta$  such that  $T_i(u' \oplus r_1, [C], \rho[a \mapsto d])$ . Since  $a$  does not occur free in  $C$ , we have  $T_i(u' \oplus r_1, [C], \rho)$ . Therefore, we are done, by Lemma 3.16 and the fact that  $u' \oplus r_1 \leq u' \oplus r$ .

$$\frac{\begin{array}{c} \vdots r_1 \\ \Gamma \rightarrow \Delta, A(t) \end{array}}{t \leq s, \Gamma \rightarrow \Delta, \exists x \leq s. A(x)} \text{Rb } \exists, \quad (3.19)$$

where the variable  $x$  does not occur in the term  $s$ .

Assume that  $T_i(u', [t \leq s], \rho)$  and  $\forall B \in \Gamma, T_i(u', [B], \rho)$ . Note that  $\rho$  is an environment for  $\Gamma_{r_1} \rightarrow \Delta_{r_1}$ . Let  $\rho_1$  be the subsequence of  $\rho$  such that  $\text{Env}(\rho_1, [\Gamma_{r_1} \rightarrow \Delta_{r_1}], u)$ . Then we have  $\forall B \in \Gamma, T_i(u', [B], \rho_1)$ , and  $\rho_1 \leq \text{BdEnv}([\Gamma_{r_1} \rightarrow \Delta_{r_1}], u)$ . By the induction hypothesis applied to  $r_1$  together with the fact that  $u' \leq u \oplus r \leq u \oplus r_1$ , either  $\exists C \in \Delta, T_i(u' \oplus r_1, [C], \rho_1)$  or  $T_i(u' \oplus r_1, [A(t)], \rho_1)$ .

If there is some  $C$  in  $\Delta$  such that  $T_i(u' \oplus r_1, [C], \rho_1)$ , then we have  $T_i(u' \oplus r_1, [C], \rho)$  by Lemma 3.17. Hence we are done by Lemma 3.16 and the fact that  $u' \oplus r_1 \leq u' \oplus r$ , so assume otherwise.

Then we have  $T_i(u' \oplus r_1, [A(t)], \rho_1)$ , hence  $T_i(u' \oplus r_1, [A(t)], \rho)$  holds by Lemma 3.17. By Lemma 3.15, there exists  $c$  such that  $v([t], \rho) \downarrow_{u' \oplus r_1} c$  and  $T_i(u' \oplus r_1, [A(a)], \rho[a \mapsto c])$ . Since  $T_i(u', [t \leq s], \rho)$  holds, there exists  $d$  such that  $v([s], \rho) \downarrow_{u'} d$  and  $c \leq d$ . By Lemma 3.18,  $T_i(u' \oplus r_1, [\exists x \leq s. A(x)], \rho)$  holds. Thus we are done, by Lemma 3.16 and the fact that  $u' \oplus r_1 \leq u' \oplus r$ .



Cut rule:

$$\frac{\begin{array}{c} \vdots r_1 \\ \Gamma \rightarrow \Delta, A \end{array} \quad \begin{array}{c} \vdots r_2 \\ A, \Pi \rightarrow \Lambda \end{array}}{\Gamma, \Pi \rightarrow \Delta, \Lambda} \text{Cut} \quad (3.20)$$

Assume that  $\forall B \in \Gamma, \Pi, T_i(u', [B], \rho)$ . Let  $\vec{a}$  be the variables that occur free in  $A$  but do not occur free in  $\Gamma_r \rightarrow \Delta_r$ , and let  $\rho[\vec{a} \mapsto \vec{0}]$  be the environment that extends  $\rho$  and maps every variable in  $\vec{a}$  to 0 (where  $\rho[\vec{a} \mapsto \vec{0}] \equiv \rho$  if  $\vec{a}$  is empty). Note that  $\rho[\vec{a} \mapsto \vec{0}]$  is an environment for  $\Gamma_{r_1} \rightarrow \Delta_{r_1}$ , and that  $\rho[\vec{a} \mapsto \vec{0}](y) \leq u$  for every variable  $y$  that occurs free in  $\Gamma_{r_1} \rightarrow \Delta_{r_1}$ . Let  $\rho_1$  be the subsequence of  $\rho[\vec{a} \mapsto \vec{0}]$  such that  $\text{Env}(\rho_1, [\Gamma_{r_1} \rightarrow \Delta_{r_1}], u)$ . Then we have  $\forall B \in \Gamma, T_i(u', [B], \rho_1)$ , and  $\rho_1 \leq \text{BdEnv}([\Gamma_{r_1} \rightarrow \Delta_{r_1}], u)$ . By the induction hypothesis applied to  $r_1$  together with the fact that  $u' \leq u \oplus r \leq u \oplus r_1$ , either  $\exists C \in \Delta, T_i(u' \oplus r_1, [C], \rho_1)$  or  $T_i(u' \oplus r_1, [A], \rho_1)$ .

If there is some  $C$  in  $\Delta$  such that  $T_i(u' \oplus r_1, [C], \rho_1)$ , then we have  $T_i(u' \oplus r_1, [C], \rho[\vec{a} \mapsto \vec{0}])$  because  $\rho_1$  is a subsequence of  $\rho[\vec{a} \mapsto \vec{0}]$  and by Lemma 3.17. Furthermore,  $T_i(u' \oplus r_1, [C], \rho)$  holds by Lemma 3.17, since none of the variables in  $\vec{a}$  occurs free in  $C$ . Thus  $T_i(u' \oplus r, [C], \rho)$  by Lemma 3.16 and the fact that  $u' \oplus r_1 \leq u' \oplus r$ . Hence we are done, so assume otherwise.

Then we have  $T_i(u' \oplus r_1, [A], \rho_1)$ , so  $T_i(u' \oplus r_1, [A], \rho[\vec{a} \mapsto \vec{0}])$  holds by Lemma 3.17, because  $\rho_1$  is a subsequence of  $\rho[\vec{a} \mapsto \vec{0}]$ . By our assumption about  $\Pi$ , we have  $\forall B \in \Pi, T_i(u', [B], \rho[\vec{a} \mapsto \vec{0}])$  by Lemma 3.17, because  $\rho$  is a subsequence of  $\rho[\vec{a} \mapsto \vec{0}]$ . By Lemma 3.16, we have  $\forall B \in \Pi, T_i(u' \oplus r_1, [B], \rho[\vec{a} \mapsto \vec{0}])$ .

Note that  $\rho[\vec{a} \mapsto \vec{0}]$  is an environment for  $\Gamma_{r_2} \rightarrow \Delta_{r_2}$ , and that  $\rho[\vec{a} \mapsto \vec{0}](y) \leq u$  for every variable  $y$  that occurs free in  $\Gamma_{r_2} \rightarrow \Delta_{r_2}$ . Let  $\rho_2$  be the subsequence of  $\rho[\vec{a} \mapsto \vec{0}]$  such that  $\text{Env}(\rho_2, [\Gamma_{r_2} \rightarrow \Delta_{r_2}], u)$ . Then we have  $T_i(u' \oplus r_1, [A], \rho_2)$  and  $\forall B \in \Pi, T_i(u' \oplus r_1, [B], \rho_2)$ ; in addition,  $\rho_2 \leq \text{BdEnv}([\Gamma_{r_2} \rightarrow \Delta_{r_2}], u)$ .

Our choice of Gödel numbering, together with the fact that  $r_1$  and  $r_2$  are Gödel numbers of nonempty subproofs of  $\Gamma_r \rightarrow \Delta_r$ , ensures that  $|r_1 \oplus r_2| < |r|$ . Since  $u' \leq u \oplus r$ , we have  $|u' \oplus r_1| \leq |u \oplus r \oplus r_1| < |u \oplus (r_1 \oplus r_2) \oplus r_1| = |u \oplus r_2|$ , hence  $u' \oplus r_1 < u \oplus r_2$ .

By the induction hypothesis applied to  $r_2$  together with the fact that  $u' \oplus r_1 < u \oplus r_2$ , there is some  $D$  in  $\Lambda$  such that  $T_i(u' \oplus r_1 \oplus r_2, [D], \rho_2)$ . Then we have  $T_i(u' \oplus r_1 \oplus r_2, [D], \rho[\vec{a} \mapsto \vec{0}])$  by Lemma 3.17, because  $\rho_2$  is a subsequence of  $\rho[\vec{a} \mapsto \vec{0}]$ . Furthermore,  $T_i(u' \oplus r_1 \oplus r_2, [D], \rho)$  holds by Lemma 3.17, because none of the variables in  $\vec{a}$  occurs free in  $D$ . Since  $|r_1 \oplus r_2| < |r|$ , we have  $u' \oplus r_1 \oplus r_2 < u' \oplus r$ , so  $T_i(u' \oplus r, [D], \rho)$  by Lemma 3.16. Hence we are done.  $\square$

**Theorem 3.21.** *Let  $i\text{-Con} \equiv \forall w. \neg i\text{-Prf}(w, [\rightarrow])$ , which states that there is no strictly  $i$ -normal proof of the empty sequent  $\rightarrow$ . Then*

$$S_2^{i+2} \vdash i\text{-Con} \quad (3.21)$$

*Proof.* We informally argue inside of  $S_2^{i+2}$ . Assume that  $i\text{-Prf}(w, [\rightarrow])$  holds for some  $w$ . Let  $u$  be as in the statement of Proposition 3.20, let  $\rho$  be the empty environment, and let  $r$  be the root of  $w$ . Then we obtain  $[\forall A \in \Gamma_r, T_i(u', [A], \rho)] \supset [\exists B \in \Delta_r, T_i(u' \oplus r, [B], \rho)]$ . However, both  $\Gamma_r$  and  $\Delta_r$  are empty. Therefore, we obtain  $[\forall A \in \emptyset, T_i(u', [A], \rho)] \supset [\exists B \in \emptyset, T_i(u' \oplus r, [B], \rho)]$ . Since there is no  $A \in \emptyset$ , the premise is true. But since there is no  $B \in \emptyset$ , the conclusion cannot be true. This is a contradiction. Therefore, the formula  $\forall w. \neg i\text{-Prf}(w, [\rightarrow])$  holds.  $\square$

#### 4. BOOTSTRAPPING THEOREM OF $S_2^i E$

In this section, we establish the correspondence between  $S_2^i E$  and  $S_2^i$ . We show that  $S_2^i E$  has essentially the same strength as  $S_2^i$  if  $i \geq 1$ . The theorem which establishes the correspondence is called the Bootstrapping Theorem (Theorem 4.2), following Buss' use of the term "bootstrapping" in [5], since we bootstrap from the restricted set of axioms of  $S_2^i E$  to the full power of  $S_2^i$ .

We present a proof of the theorem in four "phases" of bootstrapping. In the first phase, we show that all the functions of  $S_2 E$  are provably total. Each of the remaining phases applies to a particular class of inferences of  $S_2^i$ , and we show that all the inferences covered in each phase are *admissible in  $S_2^i E$*  (if properly translated from  $S_2^i$  to  $S_2^i E$ ), that is, that if all the premises of an inference covered in a given phase are provable in  $S_2^i E$ , then the conclusion of that inference is also provable in  $S_2^i E$  (Definition 4.7). The Bootstrapping Theorem (Theorem 4.2) then follows from the fact that every inference of  $S_2^i$  is treated in some phase of the bootstrapping. Even the axioms are included in this, since an axiom is just a rule of inference with no premise.

**4.1. Translation of theorems of  $S_2^i$ .** In this subsection, we introduce a translation of  $S_2^i$  formulae to the language of  $S_2^i E$  and state the Bootstrapping Theorem (Theorem 4.2).

**Definition 4.1.** The formulae of  $S_2^i$  are translated into formulae of  $S_2^i E$  by replacing every formula of the form  $A \supset B$  with one of the form  $\neg A \vee B$ , and using De Morgan duality to replace every formula of the form  $\neg A$  with a logically equivalent formula in which every subformula prefaced with the negation symbol " $\neg$ " is of the form  $t_1 = t_2$  or  $t_1 \leq t_2$ . We call this translation the *\*-translation* and denote the \*-translation of  $A$  by  $A^*$ . Formally, the \*-translation is defined as follows.

- (1)  $(p(t_1, t_2))^* \equiv p(t_1, t_2)$  if  $p$  is  $=$  or  $\leq$ .
- (2)  $(\neg p(t_1, t_2))^* \equiv \neg p(t_1, t_2)$  if  $p$  is  $=$  or  $\leq$ .
- (3)  $(A \wedge B)^* \equiv A^* \wedge B^*$ .
- (4)  $(A \vee B)^* \equiv A^* \vee B^*$ .
- (5)  $(\neg A)^* \equiv (\overline{A})^*$ , where  $\overline{A}$  is the De Morgan dual of  $A$ .
- (6)  $(A \supset B)^* \equiv (\overline{A})^* \vee B^*$ .
- (7)  $(\forall x \leq t. A)^* \equiv \forall x \leq t. A^*$  and  $(\exists x \leq t. A)^* \equiv \exists x \leq t. A^*$ .
- (8)  $(\forall x. A)^* \equiv \forall x. A^*$  and  $(\exists x. A)^* \equiv \exists x. A^*$ .

$\Gamma^*$  is the sequence of formulae which is obtained by applying  $*$  to the formulae in the sequence  $\Gamma$ .

The sequent  $\Gamma \rightarrow \Delta$  is translated to the sequent  $(\Gamma \rightarrow \Delta)^* \equiv E\vec{a}, \Gamma^* \rightarrow \Delta^*$ , where  $\vec{a}$  are the variables that occur free in  $\Gamma \rightarrow \Delta$ .

The following theorem states that  $S_2^i E$  proves the \*-translations of sequents derivable in  $S_2^i$  if  $i \geq 0$ .

**Theorem 4.2** (Bootstrapping Theorem). *If  $i \geq 1$  and  $S_2^i$  proves a sequent  $\Gamma \rightarrow \Delta$ , then  $S_2^i E(\mathcal{F}, \mathcal{A})$  proves its \*-translation  $(\Gamma \rightarrow \Delta)^*$  if  $\mathcal{F}$  and  $\mathcal{A}$  satisfy the conditions presented in Subsections 2.1 and 2.2, respectively.*

The rest of this section is devoted to a proof of the Bootstrapping Theorem. To simplify the notation, we write  $S_2^i E$  for  $S_2^i E(\mathcal{F}, \mathcal{A})$

**4.2. Bootstrapping Phase I:  $S_2^i E$  proves totality of its functions.** In this subsection, we prove that if  $i \geq 0$ , all the functions of  $S_2^i E$  are provably total, that is, that  $S_2^i E \vdash E\vec{a} \rightarrow Ef\vec{a}$  for every function symbol  $f \in \mathcal{F}$ . The proof is by induction (in the meta-language) on the definition degree of  $f$  (Definition 2.3).

**Proposition 4.3.** *If  $i \geq 0$ , then for every  $n$ -ary function symbol  $f$  of  $S_2^i E$ ,  $S_2^i E$  proves*

$$E\vec{a} \rightarrow Ef\vec{a}, \quad (4.1)$$

where  $\vec{a} \equiv a_1, \dots, a_n$ .

The reason for specifying that  $i \geq 0$  is that in the proof we apply the PIND rule to  $\Sigma_0^b$  formulae of  $S_2^i E$ .

It follows from this proposition that if all the variables in a term of  $S_2^i E$  converge, then the term itself converges.

**Corollary 4.4.** *Let  $t$  be a term of  $S_2^i E$ . If  $a_1, \dots, a_n$  are the variables that occur in  $t$ , then the following holds if  $i \geq 0$ .*

$$S_2^i E \vdash Ea_1, \dots, Ea_n \rightarrow Et \quad (4.2)$$

*Proof of Corollary 4.4.* Induction on the construction of  $t$ .

The base case  $t \equiv 0$ : This is immediate, since  $\rightarrow E0$  is Axiom (2.16).

The base case  $t \equiv a_1$ : This is also immediate, by Identity ( $Ea_1 \rightarrow Ea_1$ ).

Induction step  $t \equiv ft_1 \dots t_m$ : We assume that the corollary holds for  $t_1, \dots, t_m$ . Then for every  $j$ , we have an  $S_2^i E$  proof of  $Ea_1, \dots, Ea_n \rightarrow Et_j$  by the induction hypothesis (together with Weakening if at least one of the variables  $a_1, \dots, a_n$  does not occur in  $t_j$ ). Let  $b_1, \dots, b_m$  be variables. By Proposition 4.3, we have  $Eb_1, \dots, Eb_m \rightarrow Efb_1 \dots b_m$ . By the Substitution Lemma (Lemma 2.16), we have  $Et_1, \dots, Et_m \rightarrow Eft_1 \dots t_m$ . Applying Cut  $m$  times (once for each  $j$ ), followed by Contraction every time but the first, we obtain  $Ea_1, \dots, Ea_n \rightarrow Eft_1 \dots t_m$ .  $\square$

The rest of this subsection is devoted to a proof of Proposition 4.3.

*Proof of Proposition 4.3.* The proof is by induction on  $d(f)$ , the definition degree of the function  $f$ .

Base case:  $d(f) = 0$ . If  $d(f) = 0$ , then  $f$  is either the constant function  $0^n$ , a projection function  $\text{proj}_k^n$ , one of the binary successor functions  $s_0, s_1$ . If  $f$  is  $0^n$ , then

$$E\vec{a}, E0 \rightarrow 0^n(a_1, \dots, a_n) = 0 \quad (4.3)$$

is an axiom (See Definition 2.14). By  $\rightarrow E0$  (Axiom (2.16)) together with Cut, we have  $E\vec{a} \rightarrow 0^n(a_1, \dots, a_n) = 0$ . Using a substitution instance of Axiom (2.8) with  $p$  set to  $=$ , we can derive  $0^n(a_1, \dots, a_n) = 0 \rightarrow E0^n(a_1, \dots, a_n)$ . Thus, we obtain  $Ea_1, \dots, Ea_n \rightarrow E0^n(a_1, \dots, a_n)$  by Cut.

If  $f$  is  $\text{proj}_k^n$ , then

$$E\vec{a} \rightarrow \text{proj}_k^n(a_1, \dots, a_n) = a_k \quad (4.4)$$

is an axiom. Using a substitution instance of Axiom (2.8) with Cut, we can derive

$$Ea_1, \dots, Ea_n \rightarrow E\text{proj}_k^n(a_1, \dots, a_n). \quad (4.5)$$

If  $f$  is a binary successor function  $s_j$  for some  $j \in \{0, 1\}$ , then  $Ea \rightarrow Es_j a$  is a data axiom (Definition 2.12).

Induction step:  $f$  is defined either by recursion or by composition. We first consider the case of composition.

Assume that  $f$  is defined from functions  $g, h_1, \dots, h_m$  by composition. Then the defining axiom for  $f$  has the following form.

$$E\vec{a}, Eg(h_1(\vec{a}), \dots, h_m(\vec{a})) \rightarrow f(\vec{a}) = g(h_1(\vec{a}), \dots, h_m(\vec{a})) \quad (4.6)$$

By the induction hypothesis,  $S_2^i E$  proves

$$Eb_1, \dots, Eb_m \rightarrow Eg(b_1, \dots, b_m), \quad (4.7)$$

where  $b_1, \dots, b_m$  are variables not in  $\{a_1, \dots, a_n\}$ , and

$$Ea_1, \dots, Ea_n \rightarrow Eh_j(a_1, \dots, a_n) \quad (4.8)$$

for  $j \in \{1, \dots, m\}$ .

Substituting  $h_j(a_1, \dots, a_n)$  for  $b_j$  ( $j = 1, \dots, m$ ) in (4.7) and using the Substitution Lemma (Lemma 2.16), we obtain an  $S_2^i E$  proof of

$$Eh_1(a_1, \dots, a_n), \dots, Eh_m(a_1, \dots, a_n) \rightarrow Eg(h_1(a_1, \dots, a_n), \dots, h_m(a_1, \dots, a_n)). \quad (4.9)$$

Applying Cut to (4.9) and (4.8)  $m$  times (once for every  $j$ ), followed by Contraction every time but the first, we have an  $S_2^i E$  proof of

$$Ea_1, \dots, Ea_n \rightarrow Eg(h_1(a_1, \dots, a_n), \dots, h_m(a_1, \dots, a_n)). \quad (4.10)$$

Applying Cut to (4.10) and (4.6), followed by Contraction, we have

$$Ea_1, \dots, Ea_n \rightarrow f(\vec{a}) = g(h_1(\vec{a}), \dots, h_m(\vec{a})) \quad (4.11)$$

Finally, using a substitution instance of Axiom (2.8) together with Cut, we obtain an  $S_2^i E$  proof of

$$Ea_1, \dots, Ea_n \rightarrow Ef(a_1, \dots, a_n). \quad (4.12)$$

Next, we consider the case where  $f$  is defined from functions  $g, h_0, h_1$  by recursion. Then the defining axioms for  $f$  have the following forms.

$$E\vec{a}, Eg(\vec{a}) \rightarrow f(0, \vec{a}) = g(\vec{a}) \quad (4.13)$$

$$Ea, Eh_j(a, f(a, \vec{a}), \vec{a}) \rightarrow f(s_j a, \vec{a}) = h_j(a, f(a, \vec{a}), \vec{a}) \quad (4.14)$$

where  $j \in \{0, 1\}$ .

By the induction hypothesis,  $S_2^i E$  proves

$$Ea_1, \dots, Ea_n \rightarrow Eg(a_1, \dots, a_n) \quad (4.15)$$

and

$$Ea, Eb, Ea_1, \dots, Ea_n \rightarrow Eh_j(a, b, a_1, \dots, a_n), \quad (4.16)$$

where  $b$  is a variable not in  $\{a, a_1, \dots, a_n\}$  and  $j \in \{0, 1\}$ . Applying Cut to (4.15) and (4.13), followed by Contraction, we obtain

$$Ea_1, \dots, Ea_n \rightarrow f(0, a_1, \dots, a_n) = g(a_1, \dots, a_n) \quad (4.17)$$

Using a substitution instance of Axiom (2.8) and Cut, we can derive

$$Ea_1, \dots, Ea_n \rightarrow Ef(0, a_1, \dots, a_n) \quad (4.18)$$

Substituting  $f(a, a_1, \dots, a_n)$  for  $b$  in (4.16) and using the Substitution Lemma (Lemma 2.16), we obtain  $S_2^i E$  proofs of

$$Ea, Ef(a, a_1, \dots, a_n), Ea_1, \dots, Ea_n \rightarrow Eh_j(a, f(a, a_1, \dots, a_n), a_1, \dots, a_n) \quad (4.19)$$

for  $j \in \{0, 1\}$ .

Applying Cut to this result and (4.14), together with totality of  $\text{Cond } \ominus$ , and  $k$ , we obtain

$$Ea, Ef(a, \vec{a}), E\vec{a} \rightarrow f(s_j a, \vec{a}) = h_j(a, f(a, \vec{a}), \vec{a}) \quad (4.20)$$

for  $j \in \{0, 1\}$ . Using a substitution instance of Axiom (2.8) together with Cut, we have  $S_2^i E$  proofs of

$$Ea, Ef(a, a_1, \dots, a_n), Ea_1, \dots, Ea_n \rightarrow Ef(s_0 a, a_1, \dots, a_n) \quad (4.21)$$

$$Ea, Ef(a, a_1, \dots, a_n), Ea_1, \dots, Ea_n \rightarrow Ef(s_1 a, a_1, \dots, a_n). \quad (4.22)$$

Applying the  $\Sigma_0^b$ -PIND- $E$  rule (2.50) to (4.18), (4.21), and (4.22), and setting  $t$  to  $a$ , we have an  $S_2^i E$  proof of

$$Ea, Ea_1, \dots, Ea_n \rightarrow Ef(a, a_1, \dots, a_n). \quad (4.23)$$

This completes the induction step. □

**4.3. Bootstrapping Phase II :  $S_2^i E$  proves  $*$ -translations of axioms of  $S_2^i$ .** In Bootstrapping Phase II, we prove the  $*$ -translations of axioms of  $S_2^i$  in  $S_2^i E$ . There are two kinds of axioms: equality axioms and BASIC axioms.

**Proposition 4.5.** *The  $*$ -translations of the equality axioms of  $S_2^i$  are provable in  $S_2^i E$ .*

*Proof.* First, we consider the equality axiom  $\vec{a} = \vec{b} \rightarrow f(\vec{a}) = f(\vec{b})$ . The  $*$ -translation of this is  $E\vec{a}, E\vec{b}, \vec{a} = \vec{b} \rightarrow f(\vec{a}) = f(\vec{b})$ . By Proposition 4.3,  $S_2^i E$  proves  $E\vec{a} \rightarrow Ef(\vec{a})$ . Therefore, we can derive it from Axiom (2.14).

The proofs for the other equality axioms are straightforward. □

Next we prove that the  $*$ -translations of the BASIC axioms of  $S_2^i$  are provable in  $S_2^i E$ .

**Proposition 4.6.** *Assume that  $A$  is a BASIC axiom. Then  $(\rightarrow A)^*$  (the  $*$ -translation of  $\rightarrow A$ ) is derivable in  $S_2^i E$ .*

*Proof.* The BASIC axioms can be derived from the auxiliary axioms by Corollary 4.4 and propositional inferences. For example, we consider the BASIC axiom  $|a| = |b| \supset a \# c = b \# c$ . We have the corresponding auxiliary axiom  $|a| = |b|, Ea \# c, Eb \# c \rightarrow a \# c = b \# c$ . By Corollary 4.4, we have  $Ea, Ec \rightarrow Ea \# c$  and  $Eb, Ec \rightarrow Eb \# c$ . Therefore, we have  $|a| = |b|, Ea, Eb, Ec \rightarrow a \# c = b \# c$  by Cut and Contraction. By propositional inference,  $Ea, Eb, Ec, E|a|, E|b| \rightarrow |a| \neq |b| \vee a \# c = b \# c$ . Again, by Corollary 4.4 and Cut, we have  $Ea, Eb, Ec \rightarrow |a| \neq |b| \vee a \# c = b \# c$ , which is the  $*$ -translation of  $\rightarrow |a| = |b| \supset a \# c = b \# c$ . □

**4.4. Bootstrapping Phase III :\*-translations of predicate logic are admissible in  $S_2^i E$ .** In Bootstrapping Phase III, we prove that the \*-translations of the inferences of predicate logic are admissible in  $S_2^i E$ .

**Definition 4.7.** The inference

$$\frac{\Gamma_1 \rightarrow \Delta_1 \quad \cdots \quad \Gamma_n \rightarrow \Delta_n}{\Gamma \rightarrow \Delta} \quad (4.24)$$

is *admissible in  $S_2^i E$*  if  $\Gamma \rightarrow \Delta$  is provable in  $S_2^i E$  whenever  $\Gamma_1 \rightarrow \Delta_1, \dots, \Gamma_n \rightarrow \Delta_n$  are provable in  $S_2^i E$ .

**Proposition 4.8.** *If*

$$\frac{\Gamma_1 \rightarrow \Delta_1 \quad \cdots \quad \Gamma_n \rightarrow \Delta_n}{\Gamma \rightarrow \Delta} \quad (4.25)$$

*is an inference of predicate logic, then the inference*

$$\frac{(\Gamma_1 \rightarrow \Delta_1)^* \quad \cdots \quad (\Gamma_n \rightarrow \Delta_n)^*}{(\Gamma \rightarrow \Delta)^*} \quad (4.26)$$

*is admissible in  $S_2^i E$ .*

We prove the proposition by considering the various rules of inference of  $S_2^i$ . We begin by providing detailed proofs for the L  $\neg$ -rule and the R  $\neg$ -rule, in Lemma 4.9 and Lemma 4.10, respectively. Then we proceed to proofs for the  $\supset$ -rules and the quantifier rules. The proofs for the other rules are trivial.

**Lemma 4.9.**

$$\frac{E\vec{a}, \Gamma^* \rightarrow \Delta^*, A^*}{E\vec{a}, (\neg A)^*, \Gamma^* \rightarrow \Delta^*} \quad (4.27)$$

*is admissible in  $S_2^i E$ , where  $\vec{a}$  are the variables that occur free in the sequent  $\Gamma \rightarrow \Delta, A$ .*

*Proof.* By induction on  $A$ .

If  $A$  is atomic, the inference given in the statement of the Lemma is an instance of the L  $\neg$ -rule of  $S_2^i E$ . In the induction step, we assume that  $E\vec{a}, \Gamma^* \rightarrow \Delta^*, A^*$  holds and show that  $E\vec{a}, (\neg A)^*, \Gamma^* \rightarrow \Delta^*$  holds. The proof depends on the form of  $A$ .

$A \equiv \neg A_1$ : By Identity and Weakening, we have  $E\vec{b}, (A_1)^* \rightarrow (A_1)^*$ , where  $\vec{b}$  are the variables that occur free in  $A_1$ . By the induction hypothesis applied to  $A_1$ , we can prove  $E\vec{b}, (A_1)^*, (\neg A_1)^* \rightarrow$  in  $S_2^i E$ . By assumption,  $E\vec{a}, \Gamma^* \rightarrow \Delta^*, (\neg A_1)^*$  is derivable. Thus  $E\vec{a}, (A_1)^*, \Gamma^* \rightarrow \Delta^*$  is derivable by Cut and Contraction. By the definition of  $*$ , we are done. (Note that, by De Morgan duality,  $(\neg \neg A_1)^* \equiv (A_1)^*$ .)

$A \equiv A_1 \wedge A_2$ : Since  $(A_1 \wedge A_2)^* \rightarrow (A_1)^*$  and  $(A_1 \wedge A_2)^* \rightarrow (A_2)^*$  are derivable in  $S_2^i E$  by purely propositional reasoning,  $E\vec{a}, \Gamma^* \rightarrow \Delta^*, (A_1)^*$  and  $E\vec{a}, \Gamma^* \rightarrow \Delta^*, (A_2)^*$  are derivable from the assumption that  $E\vec{a}, \Gamma^* \rightarrow \Delta^*, (A_1)^* \wedge (A_2)^*$  is provable. By the induction hypothesis applied to  $A_1$  and  $A_2$ ,  $E\vec{a}, (\neg A_1)^*, \Gamma^* \rightarrow \Delta^*$  and  $E\vec{a}, (\neg A_2)^*, \Gamma^* \rightarrow \Delta^*$  are derivable. Thus by the L  $\vee$ -rule of  $S_2^i E$ , we can derive  $E\vec{a}, (\neg A_1)^* \vee (\neg A_2)^*, \Gamma^* \rightarrow \Delta^*$ . By the definition of  $*$ , we are done.

$A \equiv A_1 \vee A_2$ : Since  $(A_1)^* \vee (A_2)^* \rightarrow (A_1)^*, (A_2)^*$  is derivable by purely propositional reasoning in  $S_2^i E$ , we have  $E\vec{a}, \Gamma^* \rightarrow \Delta^*, (A_1)^*, (A_2)^*$  from the assumption that  $E\vec{a}, \Gamma^* \rightarrow \Delta^*, (A_1)^* \vee (A_2)^*$ . Then  $E\vec{a}, (\neg A_1)^*, \Gamma^* \rightarrow \Delta^*, (A_2)^*$ , by the induction hypothesis applied to  $A_1$ . Again applying the induction hypothesis, this time to  $A_2$ , we obtain a sequent  $E\vec{a}, \neg(A_1)^*, \neg(A_2)^*, \Gamma^* \rightarrow \Delta^*$ . By the  $L \wedge_1$ - and  $L \wedge_2$ -rules of  $S_2^i E$ , together with Contraction, we have  $E\vec{a}, \neg(A_1)^* \wedge \neg(A_2)^*, \Gamma^* \rightarrow \Delta^*$ . By the definition of  $*$ , we are done.

$A \equiv A_1 \supset A_2$ : Since  $(\neg A_1)^* \vee (A_2)^* \rightarrow (\neg A_1)^*, (A_2)^*$  is derivable by purely propositional reasoning in  $S_2^i E$ , we have  $E\vec{a}, \Gamma^* \rightarrow \Delta^*, (\neg A_1)^*, (A_2)^*$  from the assumption that  $E\vec{a}, \Gamma^* \rightarrow \Delta^*, (\neg A_1)^* \vee (A_2)^*$ . Applying the induction hypothesis twice in succession (once to  $\neg A_1$  and once to  $A_2$ ), we have  $E\vec{a}, (A_1)^*, (\neg A_2)^*, \Gamma^* \rightarrow \Delta^*$  (since  $(\neg \neg A_1)^* \equiv (A_1)^*$ ). By the  $L \wedge_1$ - and  $L \wedge_2$ -rules of  $S_2^i E$ , together with Contraction, we have  $E\vec{a}, (A_1)^* \wedge (\neg A_2)^*, \Gamma^* \rightarrow \Delta^*$ . By definition of  $*$ , we are done.

$A \equiv \forall x \leq t. A_1(x)$ : Let  $a$  be a variable that occurs in neither  $A$  nor  $\Gamma \rightarrow \Delta$ . By Identity, we have  $A_1(a)^* \rightarrow A_1(a)^*$ . Thus  $E\vec{b}, Ea, a = a, A_1(a)^* \rightarrow A_1(a)^*$  holds by Weakening, where  $\vec{b}$  are the variables other than  $a$  that occur free in  $A_1(a)$ . Since  $a$  occurs free in  $a = a, A_1(a)^*$ , we can apply the induction hypothesis to  $E\vec{b}, Ea, a = a, A_1(a)^* \rightarrow A_1(a)^*$ , hence we have  $E\vec{b}, Ea, a = a, A_1(a)^*, (\neg A_1(a))^* \rightarrow$ .

Using  $Ea \rightarrow a = a$  (Axiom (2.10)), we can derive  $E\vec{b}, Ea, A_1(a)^*, (\neg A_1(a))^* \rightarrow$  by Cut and Contraction. By the  $Lb \forall$ -rule of  $S_2^i E$ , we obtain

$$E\vec{b}, Ea, a \leq t, \forall x \leq t. A_1(x)^*, (\neg A_1(a))^* \rightarrow. \quad (4.28)$$

Using  $a \leq t \rightarrow Ea$  (a substitution instance of Axiom (2.8)), we obtain  $E\vec{b}, a \leq t, \forall x \leq t. A_1(x)^*, (\neg A_1(a))^* \rightarrow$  by Cut and Contraction. Using the  $Lb \exists$ -rule of  $S_2^i E$ , we obtain  $E\vec{b}, \forall x \leq t. A_1(x)^*, \exists x \leq t. (\neg A_1(x))^* \rightarrow$ . By assumption,  $E\vec{a}, \Gamma^* \rightarrow \Delta^*, \forall x \leq t. A_1(x)^*$  holds, so by Cut and Contraction we have  $E\vec{a}, \exists x \leq t. (\neg A_1(x))^*, \Gamma^* \rightarrow \Delta^*$ . By the definition of  $*$ , we are done.

$A \equiv \forall x. A_1(x)$ : The proof is similar to the proof of the previous case, the main differences being that we use  $Ea$  instead of  $a \leq t$  and we apply the unbounded counterparts of the bounded-quantifier rules of  $S_2^i E$  employed in that proof.

$A \equiv \exists x \leq t. A_1(x)$ : Let  $a$  be a variable that occurs in neither  $A$  nor  $\Gamma \rightarrow \Delta$ . As shown in the proof of the case where  $A \equiv \forall x \leq t. A_1(x)$ , we can derive  $E\vec{b}, Ea, A_1(a)^*, (\neg A_1(a))^* \rightarrow$ , where  $\vec{b}$  are the variables other than  $a$  that occur free in  $A_1(a)$ . By the  $Lb \forall$ -rule of  $S_2^i E$ , we obtain  $E\vec{b}, Ea, a \leq t, (A_1(a))^*, \forall x \leq t. (\neg A_1(x))^* \rightarrow$ . Using  $a \leq t \rightarrow Ea$  (a substitution instance of Axiom (2.8)), we obtain  $E\vec{b}, a \leq t, (A_1(a))^*, \forall x \leq t. (\neg A_1(x))^* \rightarrow$  by Cut and Contraction. Using the  $Lb \exists$ -rule of  $S_2^i E$ , we obtain  $E\vec{b}, \exists x \leq t. A_1(x)^*, \forall x \leq t. (\neg A_1(x))^* \rightarrow$ . By assumption,  $E\vec{a}, \Gamma^* \rightarrow \Delta^*, \exists x \leq t. A_1(x)^*$  holds, so by Cut and Contraction we have  $E\vec{a}, \forall x \leq t. (\neg A_1(x))^*, \Gamma^* \rightarrow \Delta^*$ . By the definition of  $*$ , we are done.

$A \equiv \exists x. A_1(x)$ : The proof is similar to the proof of the previous case, the main differences being that we use  $Ea$  instead of  $a \leq t$  and we apply the unbounded counterparts of the bounded-quantifier rules of  $S_2^i E$  used in that proof.

This completes the proof for the  $L \neg$ -rule.  $\square$

**Lemma 4.10.**

$$\frac{E\vec{a}, A^*, \Gamma^* \rightarrow \Delta^*}{E\vec{a}, \Gamma^* \rightarrow \Delta^*, (\neg A)^*} \quad (4.29)$$

is admissible in  $S_2^i E$ , where  $\vec{a}$  are the variables that occur free in the sequent  $A, \Gamma \rightarrow \Delta$ .

*Proof.* By induction on  $A$ .

If  $A$  is atomic ( $A \equiv p(t_1, t_2)$ , where  $p$  is  $=$  or  $\leq$ ), the inference given in the statement of the Lemma follows from the  $R \neg$ -rule of  $S_2^i E$ , together with  $E\vec{a} \rightarrow Et_1$  and  $E\vec{a} \rightarrow Et_2$ , where the latter are derivable by Corollary 4.4. In the induction step, we assume that  $E\vec{a}, A^*, \Gamma^* \rightarrow \Delta^*$  holds and show that  $E\vec{a}, \Gamma^* \rightarrow \Delta^*, (\neg A)^*$  holds. The proof depends on the form of  $A$ .

$A \equiv \neg A_1$ : By Identity and Weakening, we obtain  $E\vec{b}, (A_1)^* \rightarrow (A_1)^*$ , where  $\vec{b}$  are the variables that occur free in  $A_1$ . By the induction hypothesis applied to  $A_1$ , we have  $E\vec{b} \rightarrow (\neg A_1)^*, (A_1)^*$ . By assumption,  $E\vec{a}, (\neg A_1)^*, \Gamma^* \rightarrow \Delta^*$  is derivable, so by Cut and Contraction we have  $E\vec{a}, \Gamma^* \rightarrow \Delta^*, (A_1)^*$ . By the definition of  $*$ , we are done (note that, by De Morgan duality,  $(\neg \neg A_1)^* \equiv (A_1)^*$ ).

$A \equiv A_1 \wedge A_2$ : Since  $(A_1)^*, (A_2)^* \rightarrow (A_1)^* \wedge (A_2)^*$  is derivable, together with the assumption that  $E\vec{a}, (A_1)^* \wedge (A_2)^*, \Gamma^* \rightarrow \Delta^*$  we obtain  $E\vec{a}, (A_1)^*, (A_2)^*, \Gamma^* \rightarrow \Delta^*$ . By the induction hypothesis applied twice in succession (once to  $A_1$  and once to  $A_2$ ), we have  $E\vec{a}, \Gamma^* \rightarrow \Delta^*, (\neg A_1)^*, (\neg A_2)^*$ . By the  $R \vee_1$ - and  $R \vee_2$ -rules of  $S_2^i E$ , together with Contraction,  $E\vec{a}, \Gamma^* \rightarrow \Delta^*, (\neg A_1)^* \vee (\neg A_2)^*$  is derivable. By the definition of  $*$ , we are done.

$A \equiv A_1 \vee A_2$ : Since  $(A_1)^* \rightarrow (A_1)^* \vee (A_2)^*$  and  $(A_2)^* \rightarrow (A_1)^* \vee (A_2)^*$  are derivable, we have  $E\vec{a}, (A_1)^*, \Gamma^* \rightarrow \Delta^*$  and  $E\vec{a}, (A_2)^*, \Gamma^* \rightarrow \Delta^*$  from the assumption that  $E\vec{a}, (A_1)^* \vee (A_2)^*, \Gamma^* \rightarrow \Delta^*$ . By the induction hypothesis applied to  $A_1$  and  $A_2$ , we have  $E\vec{a}, \Gamma^* \rightarrow \Delta^*, (\neg A_1)^*$  and  $E\vec{a}, \Gamma^* \rightarrow \Delta^*, (\neg A_2)^*$ . Thus we obtain  $E\vec{a}, \Gamma^* \rightarrow \Delta^*, (\neg A_1)^* \wedge (\neg A_2)^*$  by the  $R \wedge$ -rule of  $S_2^i E$ . By the definition of  $*$ , we are done.

$A \equiv A_1 \supset A_2$ : Since  $(\neg A_1)^* \rightarrow (\neg A_1)^* \vee (A_2)^*$  and  $(A_2)^* \rightarrow (\neg A_1)^* \vee (A_2)^*$ , we obtain  $E\vec{a}, (\neg A_1)^*, \Gamma^* \rightarrow \Delta^*$  and  $E\vec{a}, (A_2)^*, \Gamma^* \rightarrow \Delta^*$  from the assumption that  $E\vec{a}, (\neg A_1)^* \vee (A_2)^*, \Gamma^* \rightarrow \Delta^*$ . By the induction hypothesis applied to  $\neg A_1$  and  $A_2$ , we have  $E\vec{a}, \Gamma^* \rightarrow \Delta^*, (A_1)^*$  (since  $(\neg \neg A_1)^* \equiv (A_1)^*$ ) and  $E\vec{a}, \Gamma^* \rightarrow \Delta^*, (\neg A_2)^*$ . By the  $R \wedge$ -rule of  $S_2^i E$ , we obtain  $E\vec{a}, \Gamma^* \rightarrow \Delta^*, (A_1)^* \wedge (\neg A_2)^*$ . By the definition of  $*$ , we are done.

$A \equiv \forall x \leq t. A_1(x)$ : Let  $a$  be a variable that occurs in neither  $A$  nor  $\Gamma \rightarrow \Delta$ . By Identity, we have  $A_1(a)^* \rightarrow A_1(a)^*$ . Thus  $E\vec{b}, Ea, a = a, A_1(a)^* \rightarrow A_1(a)^*$  holds by Weakening, where  $\vec{b}$  are the variables other than  $a$  that occur free in  $A_1(a)$ . By the induction hypothesis applied to  $A_1(a)$ , we have  $E\vec{b}, Ea, a = a \rightarrow A_1(a)^*, (\neg A_1(a))^*$ . Using  $Ea \rightarrow a = a$  (Axiom (2.10)), we obtain  $E\vec{b}, Ea \rightarrow A_1(a)^*, (\neg A_1(a))^*$  by Cut and Contraction. By the  $Rb \exists$ -rule of  $S_2^i E$ , we obtain  $E\vec{b}, Ea, a \leq t \rightarrow A_1(a)^*, \exists x \leq t. (\neg A_1(x))^*$ . Using  $a \leq t \rightarrow Ea$  (a substitution instance of Axiom (2.8)), we obtain  $E\vec{b}, a \leq t \rightarrow A_1(a)^*, \exists x \leq t. (\neg A_1(x))^*$  by Cut and Contraction. Using the  $Rb \forall$ -rule of  $S_2^i E$ , we obtain  $E\vec{b}, Et \rightarrow \forall x \leq t. A_1(x)^*, \exists x \leq t. (\neg A_1(x))^*$ . By assumption,  $E\vec{a}, \forall x \leq t. (A_1(x))^*, \Gamma^* \rightarrow \Delta^*$  holds where  $\vec{a}$  are free variables occurring in  $\forall x \leq t. (A_1(x)), \Gamma \rightarrow \Delta$ . By Cut and Contraction we have  $E\vec{a}, Et, \Gamma^* \rightarrow \Delta^*, \exists x \leq t. (\neg A_1(x))^*$ . Using  $E\vec{a} \rightarrow Et$  (Corollary 4.4), we obtain  $E\vec{a}, \Gamma^* \rightarrow \Delta^*, \exists x \leq t. (\neg A_1(x))^*$  by Cut and Contraction. By the definition of  $*$ , we are done.



$A \equiv \forall x.A_1(x)$ : The proof is similar to the proof of the previous case, the main differences being that we use  $Ea$  instead of  $a \leq t$  and we apply the unbounded counterparts of the bounded-quantifier rules of  $S_2^iE$  used in that proof.

$A \equiv \exists x \leq t.A_1(x)$ : Let  $a$  be a variable that occurs in neither  $A$  nor  $\Gamma \rightarrow \Delta$ . As shown in the proof of the case where  $A \equiv \forall x \leq t.A_1(x)$ , we can derive  $E\vec{b}, Ea \rightarrow A_1(a)^*, (\neg A_1(a))^*$ , where  $\vec{b}$  are the variables other than  $a$  that occur free in  $A_1(a)$ . By the Rb  $\exists$ -rule of  $S_2^iE$ , we obtain  $E\vec{b}, Ea, a \leq t \rightarrow \exists x \leq t.A_1(x)^*, (\neg A_1(a))^*$ . Using  $a \leq t \rightarrow Ea$  (a substitution instance of Axiom (2.8)), we obtain  $E\vec{b}, a \leq t \rightarrow \exists x \leq t.A_1(x)^*, (\neg A_1(a))^*$  by Cut and Contraction. Using the Rb  $\forall$ -rule of  $S_2^iE$ , we obtain  $E\vec{b}, Et \rightarrow \exists x \leq t.A_1(x)^*, \forall x \leq t.(\neg A_1(x))^*$ . By assumption,  $E\vec{a}, \exists x \leq t.(A_1(x))^*, \Gamma^* \rightarrow \Delta^*$  holds, so by Cut and Contraction we have  $E\vec{a}, Et, \Gamma^* \rightarrow \Delta^*, \forall x \leq t.(\neg A_1(x))^*$ . Using  $E\vec{a} \rightarrow Et$  (Corollary 4.4), we obtain  $E\vec{a}, \Gamma^* \rightarrow \Delta^*, \forall x \leq t.(\neg A_1(x))^*$  by Cut and Contraction. By the definition of  $*$ , we are done.

$A \equiv \exists x.A_1(x)$ : The proof is similar to the proof of the previous case, the main differences being that we use  $Ea$  instead of  $a \leq t$  and we apply the unbounded counterparts of the bounded-quantifier rules used in that proof.

This completes the proof for the R  $\neg$ -rule.  $\square$

*Proof of Proposition 4.8.* We prove that the  $*$ -translation of every inference of predicate logic is admissible in  $S_2^iE$ . We consider only the rules of inference for negation, implication, and quantification. The others are obvious.

L  $\neg$ -rule: This rule is treated in Lemma 4.9.

R  $\neg$ -rule: This rule is treated in Lemma 4.10.

L  $\supset$ -rule:

$$\frac{E\vec{a}, \Gamma^* \rightarrow \Delta^*, A_1^* \quad E\vec{a}, A_2^*, \Gamma^* \rightarrow \Delta^*}{E\vec{a}, (\neg A_1)^* \vee A_2^*, \Gamma^* \rightarrow \Delta^*} \quad (4.30)$$

The admissibility of this inference follows from Lemma 4.9 and the L  $\vee$ -rule of  $S_2^iE$ .

R  $\supset$ -rule:

$$\frac{E\vec{a}, (A_1)^*, \Gamma^* \rightarrow \Delta^*, (A_2)^*}{E\vec{a}, \Gamma^* \rightarrow \Delta^*, (\neg A_1)^* \vee (A_2)^*} \quad (4.31)$$

The admissibility of this inference follows from Lemma 4.10 and the R  $\vee_1$ - and R  $\vee_2$ -rules of  $S_2^iE$ , together with Contraction.

Lb  $\forall$ -rule:

$$\frac{E\vec{a}, A(t)^*, \Gamma^* \rightarrow \Delta^*}{E\vec{a}, E\vec{b}, t \leq s, \forall x \leq s.A(x)^*, \Gamma^* \rightarrow \Delta^*}, \quad (4.32)$$

where the variable  $x$  does not occur in the term  $s$ ,  $\vec{a}$  are the variables that occur free in the sequent  $A(t), \Gamma \rightarrow \Delta$ , and  $\vec{b}$  are the variables that occur in  $s$  but are not in  $\vec{a}$ . Without loss of generality, we can assume that  $x$  does not occur in  $t$ . Therefore,  $\vec{a}, \vec{b}$  are precisely the variables that occur free in  $t \leq s, \forall x \leq s.A(x), \Gamma \rightarrow \Delta$ .

From the premise, we can derive  $E\vec{a}, E\vec{b}, t \leq s, \forall x \leq s.A(x)^*, \Gamma^* \rightarrow \Delta^*$  by the Lb  $\forall$ -rule of  $S_2^iE$  and Weakening.

L $\forall$ -rule:

$$\frac{E\vec{a}', A(t)^*, \Gamma^* \rightarrow \Delta^*}{E\vec{a}, \forall x.A(x)^*, \Gamma^* \rightarrow \Delta^*}, \quad (4.33)$$

where  $\vec{a}'$  are the variables that occur free in the sequent  $A(t), \Gamma \rightarrow \Delta$  and  $\vec{a}$  are the variables that occur free in the sequent  $\forall x.A(x), \Gamma \rightarrow \Delta$ .

Clearly, every variable that occurs free in  $\forall x.A(x), \Gamma \rightarrow \Delta$  also occurs free in  $A(t), \Gamma \rightarrow \Delta$ . If there is at least one variable in  $\vec{a}'$  which is not in  $\vec{a}$ , then for each such variable  $b$ , we substitute 0 for  $b$  in both  $Eb$  and  $t$ . After repeated application of Cut with the axiom  $\rightarrow E0$ , we obtain  $E\vec{a}, A(t')^*, \Gamma^* \rightarrow \Delta^*$ , where  $t'$  is obtained by substituting 0 in  $t$  for every variable in  $\vec{a}'$  which is not in  $\vec{a}$ .

By the L $\forall$ -rule of  $S_2^i E$ , we have  $E\vec{a}, Et', \forall x.A(x)^*, \Gamma^* \rightarrow \Delta^*$ . By Corollary 4.4, we have  $E\vec{a} \rightarrow Et'$ . Hence, we obtain  $E\vec{a}, \forall x.A(x)^*, \Gamma^* \rightarrow \Delta^*$  by Cut and Contraction.

Rb $\forall$ -rule:

$$\frac{E\vec{a}, Ea, a \leq t, \Gamma^* \rightarrow \Delta^*, A(a)^*}{E\vec{a}, \Gamma^* \rightarrow \Delta^*, \forall x \leq t.A(x)^*}, \quad (4.34)$$

where neither the variable  $a$  nor the variable  $x$  occurs in the term  $t$ ;  $a$  does not occur free in  $\Gamma \rightarrow \Delta$ ; and  $\vec{a}$  are the variables other than  $a$  that occur free in the sequent  $a \leq t, \Gamma \rightarrow \Delta, A(a)$ . Clearly,  $\vec{a}$  are precisely the variables that occur free in the sequent  $\Gamma \rightarrow \Delta, \forall x \leq t.A(x)$ .

Using  $a \leq t \rightarrow Ea$  (a substitution instance of Axiom (2.8)), we can eliminate the  $Ea$  in the antecedent of the premise by Cut and Contraction. Therefore, we have  $E\vec{a}, a \leq t, \Gamma^* \rightarrow \Delta^*, A(a)^*$ . Using the Rb $\forall$ -rule of  $S_2^i E$ , we can derive  $E\vec{a}, Et, \Gamma^* \rightarrow \Delta^*, \forall x \leq t.A(x)^*$ . However, since  $Et$  is derivable from  $E\vec{a}$ , we obtain  $E\vec{a}, \Gamma^* \rightarrow \Delta^*, \forall x \leq t.A(x)^*$ .

R $\forall$ -rule:

$$\frac{E\vec{a}\{Ea\}, \Gamma^* \rightarrow \Delta^*, A(a)^*}{E\vec{a}, \Gamma^* \rightarrow \Delta^*, \forall x.A(x)^*}, \quad (4.35)$$

where the variable  $a$  does not occur free in  $\Gamma \rightarrow \Delta$ , and  $\vec{a}$  are the variables other than  $a$  that occur free in the sequent  $\Gamma \rightarrow \Delta, A(a)$ . Clearly,  $\vec{a}$  are precisely the variables that occur free in the sequent  $\Gamma \rightarrow \Delta, \forall x.A(x)$ . Here, the formula  $Ea$  is enclosed in braces to indicate that it is not included in the premise of the  $*$ -translation of the R $\forall$ -rule unless the variable  $a$  occurs free in  $A(a)$ . Otherwise,  $Ea$  can be added to the premise by Weakening. In either case, we can derive  $E\vec{a}, Ea, \Gamma^* \rightarrow \Delta^*, \forall x.A(x)^*$  by the R $\forall$ -rule of  $S_2^i E$ .

$\exists$ -rules: The proofs of admissibility of the  $*$ -translations of the  $\exists$ -rules are analogous to the proofs of admissibility of the  $*$ -translations of the  $\forall$ -rules.  $\square$

**4.5. Bootstrapping Phase IV :  $*$ -translation of  $\Sigma_i^b$ -PIND rule is admissible in  $S_2^i E$ .** Finally, we prove admissibility of the  $*$ -translation of the  $\Sigma_i^b$ -PIND rule of  $S_2^i$ .

First, we prove that our formulation of PIND, which uses the binary successor functions, proves Buss' formulation of PIND [5], which uses  $\lfloor x/2 \rfloor$ .

**Lemma 4.11.** *Assume that  $\Gamma, Ea, A(\lfloor \frac{1}{2}a \rfloor) \rightarrow A(a), \Delta$  is provable in  $S_2^i E$ , where the variable  $a$  does not occur free in  $\Gamma \rightarrow \Delta$  and  $A(a)$  is a  $\Sigma_i^b$  formula. Then  $\Gamma, E\vec{a}, A(0) \rightarrow A(t), \Delta$  is also provable in  $S_2^i E$ , where  $\vec{a}$  are the variables that occur in the term  $t$ .*

*Proof.* Note that  $\lfloor \frac{1}{2}s_0a \rfloor = \lfloor \frac{1}{2}s_1a \rfloor = a$  if  $Ea$  holds. Therefore, substituting  $s_0a$  and  $s_1a$  for  $a$  in  $\Gamma, Ea, A(\lfloor \frac{1}{2}a \rfloor) \rightarrow A(a), \Delta$  and applying Cut with  $Ea \rightarrow Es_0a$  and  $Ea \rightarrow Es_1a$ , we obtain  $\Gamma, Ea, A(a) \rightarrow A(s_0a), \Delta$  and  $\Gamma, Ea, A(a) \rightarrow A(s_1a), \Delta$ , respectively. Combining  $\Gamma, A(0) \rightarrow A(0), \Delta$  and the  $\Sigma_i^b$ -PIND- $E$  rule, we have  $\Gamma, Et, A(0) \rightarrow A(t), \Delta$ . Since  $Et$  is derivable from  $E\vec{a}$  (Corollary 4.4), we have  $\Gamma, E\vec{a}, A(0) \rightarrow A(t), \Delta$ .  $\square$

**Proposition 4.12.** *The  $*$ -translation of the PIND rule of  $S_2^i$ , that is, the inference*

$$\frac{E\vec{a}\{, Ea\}, \Gamma^*, A(\lfloor a/2 \rfloor)^* \rightarrow A(a)^*, \Delta^*}{E\vec{a}\{, E\vec{b}\}, \Gamma^*, A(0)^* \rightarrow A(t)^*, \Delta^*}, \quad (4.36)$$

*is admissible in  $S_2^iE$ , where the variable  $a$  does not occur free in  $\Gamma \rightarrow \Delta$ ,  $A(a)$  is a  $\Sigma_i^b$  formula,  $\vec{a}$  are the variables other than  $a$  that occur free in  $\Gamma, A(\lfloor a/2 \rfloor) \rightarrow A(a), \Delta$ , and  $\vec{b}$  are the variables that occur in  $t$  but are not in  $\vec{a}$ .*

*The formula  $Ea$  (in the antecedent of  $E\vec{a}\{, Ea\}, \Gamma^*, A(\lfloor a/2 \rfloor)^* \rightarrow A(a)^*, \Delta^*$ ) is enclosed in braces, as is  $E\vec{b}$  (in the antecedent of  $E\vec{a}\{, E\vec{b}\}, \Gamma^*, A(0)^* \rightarrow A(t)^*, \Delta^*$ ), to indicate that  $Ea$  and  $E\vec{b}$  are not included in those antecedents unless the variable  $a$  occurs free in  $A(a)$ .*

*Proof.* If  $a$  does not occur free in  $A(a)$ , then the premise and the conclusion of (4.36) are identical hence (4.36) is admissible.

If  $a$  occurs free in  $A(a)$ , then since  $A(a)^*$  is a  $\Sigma_i^b$  formula, we can apply Lemma 4.11 to obtain  $E\vec{a}, E\vec{b}, \Gamma^*, A(0)^* \rightarrow A(t)^*, \Delta^*$ .  $\square$

Finally, we have the tools to prove Theorem 4.2, which was the main objective of this section.

*Proof of Theorem 4.2.* By Propositions 4.5 and 4.6, the  $*$ -translations of the axioms of  $S_2^i$  are provable in  $S_2^iE$ . Furthermore, Propositions 4.8 and 4.12 guarantee admissibility of the  $*$ -translations of the inferences of  $S_2^i$ . Therefore, the  $*$ -translation of any sequent  $\Gamma \rightarrow \Delta$  which is provable in  $S_2^i$  is provable in  $S_2^iE$ .

In addition, the proofs of these propositions show that if a proof of an  $S_2^i$  sequent  $\Gamma \rightarrow \Delta$  contains only  $\Sigma_i^b$  and  $\Pi_i^b$  formulae of  $S_2^i$ , then there is a proof of  $(\Gamma \rightarrow \Delta)^*$  that contains only  $\Sigma_i^b$  and  $\Pi_i^b$  formulae of  $S_2^iE$ .  $\square$

## 5. FINITISTIC GÖDEL SENTENCES OF $S_2^{-1}E$

In this section, we investigate finitistic Gödel sentences of  $S_2^{-1}E$ . Throughout this section,  $i$  denotes a positive integer, and  $\lceil A(\underline{x}) \rceil$  denotes the Gödel number of the formula obtained from  $A(x_1)$  by substituting for the variable  $x_1$  the numeral representation of the natural number  $x$ . The purpose of underscoring the  $x$  in  $\lceil A(\underline{x}) \rceil$  is to indicate that the entity which is substituted is the numeral *representation* of the natural number  $x$  (in the meta-language), and not the *variable*  $x$  (in the object language).

Now, let us motivate our investigation. The most interesting question concerning  $S_2^{-1}E$  is whether  $S_2^i$  proves  $i$ -Con or not. If the answer is negative, we have  $S_2^i \neq S_2^{i+2}$ , hence we can conclude that the hierarchy  $S_2^1, S_2^2, S_2^3, \dots$  does not collapse.

Buss and Ignjatović [6] used a finitistic Gödel sentence, together with Solovay's induction speed-up method, to show that  $S_2^i$  does not prove the consistency of proofs that are comprised entirely of  $\Sigma_i^b$  and  $\Pi_i^b$  formulae and use only BASIC axioms and the rules of inference of predicate logic. Since their notion of proofs and  $i$ -normal proofs have a certain

similarity, it looks as though it would be worthwhile to emulate their method to prove  $S_2^i \not\vdash i\text{-Con}$ . Unfortunately, the induction speed-up method in the form in which Buss and Ignjatović used it does not work for  $S_2^{-1}E$ . However, it would still be interesting to investigate these Gödel sentences.

First, let us see why the induction speed-up method does not work. Buss and Ignjatović employed two approaches to the induction speed-up method. In their first approach, they bounded (provably in  $S_2^1$ ) the size of an  $S_2^{-1}$  proof of a formula  $\phi(\underline{x})$  by  $(x\#(x\#x))^m + n$ , where  $m$  and  $n$  depend on the size of an  $S_2^i$  proof of  $\forall x.\phi(x)$ . In their other approach, which they applied to PV and  $PV^-$  (the induction-free fragment of PV), they presented a polynomial-time computable function which converts a PV proof of a sequent  $\Gamma \rightarrow \Delta$  with *numerically restricted variables* (see p. 241 of [6]) to a  $PV^-$  proof of that sequent. We will use Proposition 3.20 to show that neither of these approaches works for  $S_2^{-1}E$ .

Consider the first approach. If it does work for  $S_2^iE$ , then for every  $i$ -normal formula  $\phi(x)$  that satisfies  $S_2^iE \vdash \forall x.\phi(x)$ , there are natural numbers  $m, n$  and  $k$  that satisfy the condition

$$S_2^1 \vdash \forall x \exists w \leq t_k(x)^m + n. i\text{-Prf}(w, \lceil \phi(\underline{x}) \rceil), \quad (5.1)$$

where  $t_k(x) \equiv \overbrace{x\#\dots\#x}^k$ . Let  $l > k$ , and derive a contradiction by considering strictly  $i$ -normal proofs of  $t_l(a) = t_l(a)$ . By (5.1), for every natural number  $d$  there exists a strictly 1-normal (hence a strictly  $i$ -normal) proof  $w(d)$  of the formula  $t_l(d) = t_l(d)$  such that  $w(d) \leq t_k(d)^m + n$ . By Proposition 3.20,  $T_i(w(d), \lceil t_l(d) = t_l(d) \rceil, \rho)$  holds, where  $\rho$  is the empty environment. By Lemma 3.11,  $w(d) > t_l(d)$ . However,  $t_l(d) > t_k(d)^m + n$  for sufficiently large  $d$ . This contradicts (5.1).

Now, consider the second approach. Assume that it does work for  $S_2^1E$ , that is, that there is a polynomial-time computable function  $f$  which converts  $S_2^1E$  proofs of numerically restricted sequents to strictly 1-normal proofs of those sequents. Let  $t_k(x)$  as in (5.1), and consider the sequent  $\rightarrow Et_k(d)$  where  $d$  is a numeral. Then  $S_2^1E \vdash \rightarrow Et_k(d)$ , by Proposition 4.3 and the fact that  $S_2^1E \vdash Ed$ . The proof of  $Et_k(d)$  can be taken to be of size  $O(|d| \cdot 2^k)$ . Obviously,  $\rightarrow Et_k(d)$  is a sequent with numerically restricted variables (since it has no variable). Since induction speed-up works,  $f$  computes a strictly 1-normal (hence a strictly  $i$ -normal) proof  $w(d, k)$  of  $\rightarrow Et_k(d)$  in polynomial time. Hence the size of  $w(d, k)$  is less than  $O(|d|^c \cdot 2^{k \cdot c})$ . By Proposition 3.20  $T(w(d, k), \lceil Et_k(d) \rceil, \rho)$  holds, where  $\rho$  is the empty environment. By Lemma 3.11,  $w(d, k) \geq \rho(t(d))$ . Therefore, the size of  $w(d, k)$  is greater than  $O(|d|^k)$ . However, for enough large  $k$  and  $d$ ,  $O(|d|^c \cdot 2^{k \cdot c}) < O(|d|^k)$ . Contradiction.

Next, we investigate a countably infinite set of Gödel sentences  $\forall x.\varphi_k(x)$  and the terms

$t_k(x) \equiv \overbrace{x\#(x\#(\dots(x\#x)))}^k$ . We show that  $i\text{-Con}$  is equivalent to  $\forall x.\varphi_k(x)$  for every  $k \geq 2$ . Therefore, we could show that  $S_2^i \not\vdash i\text{-Con}$  by showing that  $S_2^i \not\vdash \forall x.\varphi_k(x)$  for some  $k \geq 2$ .

For  $k \geq 1$ , let  $\varphi_k$  be a formula which satisfies

$$S_2^1 \vdash \forall x [\varphi_k(x) \leftrightarrow \neg \exists w \leq t_k(x). i\text{-Prf}(w, \lceil \varphi_k(\underline{x}) \rceil)]. \quad (5.2)$$

We would like to prove that the Gödel sentence  $\forall x.\varphi_k(x)$  is undecidable. We can easily see that  $S_2^i \not\vdash \neg \forall x.\varphi_k(x)$ . For suppose that  $S_2^i \vdash \neg \forall x.\varphi_k(x)$  then  $\forall x.\varphi_k(x)$  is false by soundness of  $S_2^i$ , so there exists  $d$  such that  $\varphi_k(d)$  is false. By (5.2),  $\exists w \leq t_k(d). i\text{-Prf}(w, \lceil \varphi_k(d) \rceil)$  is true. Hence  $\varphi_k(d)$  has a strictly  $i$ -normal proof. In particular,  $S_2^i \vdash \varphi_k(d)$ . Since  $\varphi_k(d)$  is false, this contradicts the soundness of  $S_2^iE$ .

On the other hand, it looks as though it would be difficult to prove that  $S_2^i \not\vdash \forall x.\varphi_k(x)$ . The crux of the problem is that even if  $S_2^i \vdash \forall x.\varphi_k(x)$ , there is no (known) bound on the length of a strictly  $i$ -normal proof  $w(d)$  of  $\varphi_k(d)$ . We know that  $w(d)$  is not bounded by  $t_k(d)$ , because by (5.2) that would contradict the consistency of  $S_2^i$ . However, perhaps  $w(d)$  is bounded by  $t_l(d)$  for some  $l > k$ .

Finally, we prove that  $S_2^i \vdash i\text{-Con} \leftrightarrow \forall x.\varphi_k(x)$  for  $k \geq 2$ . The proof of  $S_2^i \vdash i\text{-Con} \rightarrow \forall x.\varphi_k(x)$  uses a method similar to that used in the proof of Theorem 4 on p. 135 of [5]. The proof of  $S_2^i \vdash i\text{-Con} \leftarrow \forall x.\varphi_k(x)$  is a consequence of the trivial fact that a contradiction proves anything.

First, consider  $S_2^i \vdash i\text{-Con} \rightarrow \forall x.\varphi_k(x)$ . By (5.2), we have  $S_2^1 \vdash \neg\varphi_k(x) \rightarrow \exists w \leq t_k(x).i\text{-Prf}(w, \lceil \varphi_k(\underline{x}) \rceil)$ . On the other hand, using a method similar to that used for the proof of Theorem 4 on p. 135 of [5], we can prove that for every  $\Sigma_1^b$  formula  $\psi(x)$ , there is a term  $u(x)$  such that  $S_2^1 \vdash \psi(x) \rightarrow \exists w \leq u(x).0\text{-Prf}(w, \lceil \psi(\underline{x}) \rceil)$ . Since  $\neg\varphi_k(x)$  is a  $\Sigma_1^b$  formula, there is a term  $v(x)$  such that  $S_2^1 \vdash \neg\varphi_k(x) \rightarrow \exists w \leq v(x).0\text{-Prf}(w, \lceil \neg\varphi_k(\underline{x}) \rceil)$ . Therefore,  $S_2^1 \vdash \neg\varphi_k(x) \rightarrow \exists w.i\text{-Prf}(w, \lceil \neg \rceil)$ .

Now consider  $S_2^i \vdash \forall x.\varphi_k(x) \rightarrow i\text{-Con}$ , and assume that  $i\text{-Prf}(w, \lceil \neg \rceil)$ . Then by Weakening there exists  $w'(x) = O(x)$  such that  $i\text{-Prf}(w'(x), \lceil \varphi_k(\underline{x}) \rceil)$ . Therefore,  $w'(x) \leq t_k(x)$  for sufficiently large  $x$  if  $k \geq 2$ . (Note that for  $k = 1$ , we have  $t_k(x) = x$ , so this inequality does not hold.) Hence  $\exists w' \leq t_k(x).i\text{-Prf}(w', \lceil \varphi_k(\underline{x}) \rceil)$  for sufficiently large  $x$ , in which case  $S_2^i \vdash \neg\varphi_k(x)$  by (5.2). From this it follows that  $S_2^i \vdash \exists x.\neg\varphi_k(x)$ .

#### ACKNOWLEDGEMENT

I thank Georgia Martin for thoroughly checking the drafts of this paper and proposing many improvements.

#### REFERENCES

- [1] Z Adamowicz. Herbrand consistency and bounded arithmetic. *Fundamenta Mathematicae*, 2:1–15, 2002.
- [2] Z Adamowicz and K. Zdanowski. Lower bounds for the provability of Herbrand consistency in weak arithmetics. *Fundamenta Mathematicae*, 212(3):191–216, 2011.
- [3] Zofia Adamowicz and Pawe Zbierski. On Herbrand consistency in weak arithmetic. *Archive for Mathematical Logic*, 40(6):399–413, August 2001.
- [4] Arnold Beckmann. Proving consistency of equational theories in bounded arithmetic. *Journal of Symbolic Logic*, 67(1):279–296, mar 2002.
- [5] Samuel R. Buss. *Bounded arithmetic*. Bibliopolis, 1986.
- [6] Samuel R. Buss and Aleksandar Ignjatović. Unprovability of consistency statements in fragments of bounded arithmetic. *Annals of pure and applied logic*, 74:221–244, 1995.
- [7] Petr Hájek and Pavel Pudlák. *Metamathematics of first-order arithmetic*. Springer-Verlag, 1998.
- [8] J Krajíček, Pavel Pudlák, and Gaisi Takeuti. Bounded arithmetic and the polynomial hierarchy. *Annals of pure and applied logic*, 52(1-2):143–153, April 1991.
- [9] J Krajíček and Gaisi Takeuti. On induction-free provability. *Annals of mathematics and artificial intelligence*, 6:107–125, 1992.
- [10] Daniel Leivant. Termination proofs and complexity certification. In Naoki Kobayashi and Benjamin C. Pierce, editors, *Theoretical Aspects of Computer Software*, volume 2215 of *Lecture notes in computer science*, pages 183–200. Springer Berlin / Heidelberg, 2001.
- [11] John Nolt. Free logic (Stanford encyclopedia of philosophy). <http://plato.stanford.edu/entries/logic-free/>, April 2010.
- [12] R. Parikh. Existence and feasibility in arithmetic. *Journal of Symbolic Logic*, 36(3):494–508, 1971.

- [13] P. Pudlák. Cuts, consistency statements and interpretations. *The Journal of Symbolic Logic*, 50(2):423–441, 1985.
- [14] P. Pudlák. A note on bounded arithmetic. *Fundamenta mathematicae*, 136:85–89, 1990.
- [15] Dana Scott. Identity and existence in intuitionistic logic. In Michael Fourman, Christopher Mulvey, and Dana Scott, editors, *Applications of sheaves*, volume 753 of *Lecture notes in mathematics*, pages 660–696. Springer Berlin / Heidelberg, 1979.
- [16] Gaisi Takeuti. Bounded arithmetic and truth definition. *Annals of pure and applied logic*, 39:75–104, 1988.
- [17] Gaisi Takeuti. Incompleteness theorems and  $S_2^i$  versus  $S_2^{i+1}$ . In *Logic Colloquium '96: Proceedings of the Colloquium held in San Sebastián, Spain, July 9-15, 1996*, volume 12 of *Lecture notes in logic*, pages 247–261, 1996.
- [18] Gaisi Takeuti. Gödel sentences of bounded arithmetic. *Journal of Symbolic Logic*, 65(3):1338–1346, September 2000.
- [19] A. Wilkie and J. Paris. On the scheme of induction for bounded arithmetic formulas. *Annals of pure and applied logic*, 35:261–302, 1987.